[*Translating and the computer 10*. Proceedings of a conference … 10-11 November 1988, ed. Pamela Mayorcas (London: Aslib, 1990)]

Themes in the work of Margaret Masterman

Yorick Wilks

Computing Research Laboratory, New Mexico State University, New Mexico, USA

INTRODUCTION

In this paper I shall seek to reintroduce and then focus the work of Margaret Masterman by enumerating a number of themes in her work. Some of these have been successful – in the sense of appearing, usually rediscovered, in some established place in the field of natural language processing, while others, it must be said, appear to have failed, even though they remain highly interesting. This last is a dangerous claim of course, one that can be reversed at any time. There is in my view a third category, of general programmes rather than particular representational methods, as to which one can only say that they remain unproved. In spite of their breadth, scope and originality it must also be conceded that Margaret Masterman did not have theories to cover all aspects of what would be considered the core issues of computational linguistics today: for example, she had little or nothing to say on what would now be called text theory or pragmatics. Nor did she have any particular reason for ignoring them, other than that she thought the problems that she chose to work on were in some sense the most fundamental. I shall continue here the affectionate tradition of referring to her as MMB, the initials of her married name, Margaret Masterman Braithwaite.

IDEOGRAMS

Ideograms were an early interest of MMB's (Masterman, 1954) that persisted throughout her intellectual life: the notion that ideograms were

a fundamental form of language and were of non-arbitrary interpretation. The root of this idea lay in Wittgenstein's interest (1922) in how pictures could communicate: in how the drawing of an arrow could convey movement or pointing and, before that, in his so-called Picture Theory of Truth, where objects were arranged to express facts.

The connection of all this to ideograms had been noted by Richards, who was much preoccupied by Chinese, and who developed *English through pictures* (Richards and Gibson, 1952), a highly successful language teaching tool. MMB came to Chinese through Michael Halliday, then a Cambridge University lecturer in Chinese, and began to use stick-pictures as representations of situations but which could also provide a plausible referential underpinning for language: something universal, and outside the world of the language signs themselves, yet which did not fall back on the naive referentialism of those who said that the meanings of words were things or inexpressible concepts.

Frege (new translation, 1960) had tackled this issue long before and created a notation in which propositions had a sense, but could only refer to the true or the false (at which point all differences between them, except truth value, were lost). This reference to situations, that MMB helped keep alive, has found formal expression again in Barwise and Perry's *Situation semantics* (1983). They, too, advocate a central notion of a situation as being what an utterance points to, and they too resort to cartoon-like pictures but, unlike MMB, nowhere acknowledge the role of Wittgenstein's Picture Theory of Truth.

It is as hard to capture the future in this field as in any other, and the movement of a (partially) ideogrammatical language like Japanese to the centre of the language-processing stage may yet show the importance of ideograms for grasping the nature of language. But it must be said that the relevance of ideograms for the languages of the West has yet to be shown.

LATTICES AND FANS

Although not a formalist herself, and considered an anti-formalist by many, MMB nevertheless believed passionately in the applicability of mathematical techniques to natural language; without them, she believed, there would be nothing worthy of the name of theory. Her opposition was to the assumption that formal logic, in particular, applied directly to natural language, and she would not concede much distinction between that and the methods of Chomsky (1965), a position that has some historical justification.

The two structures from which she hoped the most were lattices and 'fans', a notion she derived from some work of Brouwer (1952). MMB believed lattices (Masterman, 1959a) to be the underlying structure of thesauri while fans (Masterman, 1957a) mapped the spreading out of the

new senses of words, indefinitely into the future. She spent some time trying to amalgamate both representations into a single structure. These efforts have not met with much success nor have they been taken up by others, although Zellig Harris did at one time toy with lattices as language structures, and Mellish has recently (1988) sought to link lattice structures again to Halliday's categories of grammar and semantics.

Another problem is that fans are too simple to capture much: they have no recursive structure. Lattices are so restrictive: once it is conceded that neither words nor things fall neatly under a taxonomic tree structure, it by no means follows that they fall under a graph as tidy as a lattice either. More promising routes have been found through general applications of the theory of graphs where the constraints on possible structures can be determined empirically rather than *a priori.*

THESAURI AND THE USE OF LARGE-SCALE LANGUAGE RESOURCES

MMB believed 30 years ago that constructed entities such as dictionaries and thesauri (especially the latter) constituted real resources for computational language processing (Masterman, 1956, 1959b). That was at a time when any computational operations on such entities were often dismissed by those working in other areas of computational linguistics as low-grade concordance work. Betty May compacted the whole of *Roget's thesaurus* for MMB, from 1,000 'heads' to 800, and had them card-punched. That formed the basis for a range of experiments on Hollerith sorting machines which contributed to Karen Sparck Jones' seminal thesis work *Synonymy and semantics classification* (1964, 1986). MMB believed that thesauri such as Roget's were not just fallible human constructs but real resources with some mathematical structure that was also a guide to the structures which humans use to 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 underlying generalisations and patterns.

In recent years there has been some revival of interest in computational lexicography that has fulfilled some of MMB's hopes and dreams. It has been driven to some extent by the availability from publishers of machine-readable English dictionaries, such as *Longman's dictionary of contemporary English* (LDOCE) and *Collins-Birmingham University International Language Database* (COBUILD), with definitions written in a semi-formal way. This makes it much easier for a computational parser to extract information from them. But the initial work in the current wave was done by Amsler (1980) at Texas using *Webster's,* an old-fashioned dinosaur of a dictionary. He developed a notion of 'tangled

hierarchies' which captures the notion MMB promoted so as to get away from straightforward tree-like hierarchies.

Current centres for such work include Cambridge, Bellcore, IBM (New York), Waterloo and New Mexico, where a number of techniques have been developed, including searching for taxonomic structures, by parsing the English of the dictionary entries, and by collocational techniques applied to the word occurrences in the entries themselves. This last normally involves the construction in a computer of very large matrices, as foreseen in the earlier work of Sparck Jones. Those matrices can now be computed effectively with modern machines in a way that was virtually impossible 25 years ago.

Although dictionaries and thesauri are in some senses inverses of each other, they also differ importantly in that dictionaries are written in words that are themselves sense-ambiguous, except, that is, for those entries in a dictionary which are written as lists of semi-synonyms (as when, for example, 'gorse' is defined as 'furze'). One of the major barriers to the use of machine-readable dictionaries has been the need to resolve those lexical ambiguities as the dictionary itself is parsed, which is to say, transformed by computer into some more formal, tractable structure. MMB was more concerned with thesauri than dictionaries as practical and intellectual tools, and they do not suffer from the problem in the same way. Words in a thesaurus are also ambiguous items, but their method of placement determines their sense in a clearer way than in a dictionary: the item 'car', for example, appears in a thesaurus in a list of vehicles, and therefore means a vehicle, not a Lisp (list processing) head. The name 'vehicle', at the head of the section, can thus straightforwardly determine the sense of items in it.

THE USE OF INTERLINGUAS

MMB was much associated with the use of interlinguas for MT (machine translation) and for meaning representation (Masterman, 1967), and her reply to Bar-Hillel's criticism of their use has been much quoted. The notion of a uniform and universal meaning representation for translating between languages has continued to be a strategy within the field: it had a significant role in AI (artificial intelligence) systems such as conceptual dependency (Schank, 1975) and preference semantics (Wilks, 1973), and is now to be found in recent attempts to use Esperanto as an interlingua for MT.

MMB's own view was heavily influenced by the interlingua NUDE (for naked ideas or the bare essentials of language) first created by R. H. Richens at Cambridge for plant biology: in a revised form it became the interlingua with which CLRU (Cambridge Language Research Unit) experimented. NUDE had recursively-constructed bracketed formulas

made up from an inventory of semantic primitives, and the formulas expressed the meaning of word senses in English. Karen Sparck Jones worked on making NUDE formulas less informal and one of my own earliest efforts was to define the syntactic form of those entries, so that a revised form of NUDE became my representational system for some years. In that system the function of some of Richens' more 'prepositional' primitives were merged with what were later to become case labels, in the sense of Fillmore's Case Grammar (e.g. TO functioned very like the Destination Case) (1968).

However, MMB's attitude to these primitives was very unlike that of other peddlers of conceptual primitives or languages of thought: at no point did she suggest, in the way that became fashionable later in cognitive science, that the primitive names constituted some sort of language in the mind or brain (Fodor's view, 1975) or that, although they appeared to be English, the primitives like 'Move' and 'Do' were really the names of underlying entities that were not in any particular language at all. This kind of naive imperialism of English has been the bane of linguistics for many years, and shows, by contrast, the far greater sophistication of the structuralism that preceded it.

MMB was far too much the Wittgensteinian for such a defence of primitive entities, in this as in other matters: one could make up tiny toy languages to one's heart's content (and NUDE was exactly a toy language of 100 words) but one must never take one's language game totally seriously (linguists forget this rule). So, for her, NUDE remained a language, with all the features of a natural one like English or French, including the extensibility of sense already discussed. That tactic avoided all the problems of exactly how to justify the items and structure of a language claimed to be universal, or brain-embedded, of course, but produced its own problems such as that of what has been achieved by reducing one natural language to another, albeit a smaller and more regular one. This, of course, is exactly the question asked of the group who propose using Esperanto as the pivot for MT. MMB would put such questions forcefully to those in CLRU who showed any sign of actually believing in NUDE as having any special properties over and above those of ordinary languages, a possibility she had herself certainly entertained: this was a technique of cultural revolution, known to Zen Buddhists, and late perfected by Mao Tse Tung.

At bottom, she believed that such interlinguas were in need of some form of empirical justification and could not be treated as unprovable and arbitrary assumptions for a system, in the way Katz (1972) has tried to do by arguing from the role of assumed entities in physics and mathematics. There was one weak form of empirical support available: statistics derived from dictionaries showed that the first 100 commonest defining words in English dictionaries (exempting 'a' and 'the') corresponded very

closely indeed to the primitives of NUDE. But MMB wanted something more structural and spent some considerable time trying to associate the NUDE elements with the classifying principles of the thesaurus itself, which would then link back to the distributional facts about texts that the thesaurus itself represented. In this, as in other ways, MMB had more intuitive sympathy with earlier distributional or structural linguistics than with the more apparently mathematical and symbolic linguistics of Chomsky and his followers.

THE CENTRALITY OF MACHINE TRANSLATION AS A TASK

There is no doubt that MT, in recent years, has become a solvable task, at least for some well-specified needs, sometimes by the use of new representational theories, but more usually by means of better software engineering techniques applied to the old methods. Merely doing that has yielded better results than could have been dreamed of two decades ago.

MMB must be credited with helping to keep belief in MT alive during long years of public scepticism, and above all with the belief that MT was an intellectually challenging and interesting task (Masterman, 1957b, 1961). I think that is now widely granted, although it was not conceded within artificial intelligence, for example, until relatively recently. There it was generally thought that, although language understanding in general required inference knowledge of the world and processing of almost arbitrary complexity, MT did not: it was a task that required only superficial processing of language. I think that now almost everyone concedes that that view is false.

What MMB sought was a compromise system of meaning representation for MT: one that was fundamental to the process of translation, but did not constitute a detailed representation of all the relevant knowledge of the world. She believed there was a level of representation, linguistic if you will, probably vague as well, but which was sufficient for MT. In that sense, she totally denied the assumption behind Bar-Hillel's critique of MT (1953) – which was taken up by some artificial intelligence researchers afterwards (although not, of course, the same ones as referred to in the last paragraph) – that MT and language understanding in general did require the explicit representation of all world knowledge. This position of hers cannot be separated from her quasi-idealist belief (see further below) that world knowledge cannot be represented independently of *some* language, and hence that any true distinction between meaning representation and the representation of world knowledge is, ultimately, misconceived (see her discussion of Whorf [Masterman, 1961]). The only dispute can be about the 'level' or 'grain' of representation that particular acts of translation require.

PARSING TEXT BY SEMANTIC METHODS

A major concern of MMB's was always how to parse written English into a machine representation for MT (Masterman, 1968). She believed that such a representation should be fundamentally semantic in nature (i.e. based on meaning rather than syntax) and that those semantic structures should be used in the parsing process itself. The latter view was highly original, since virtually no one has ever proposed such a thing – the doctrine is now known as semantic parsing, and is well known even if not as fashionable as it was ten years ago – and espousing it certainly set MMB apart from the prevailing syntactic approaches of her time. Some contemporary clarification will be needed in later commentary on this point, since the meaning of the word 'semantics' as used by MMB in this connection, cannot be equated with either its use in 'semantic grammar' (e.g. Burton, 1978) to mean parsing by the use of particular word-names as they occur in text (e.g. as in a program that knew what words would follow 'electrical'), nor with its currently dominant use in formal, logical semantics, to which we shall return in a moment.

One of MMB's main motivations for her view was that natural languages are highly ambiguous as to word sense, and that this fact had been systematically ignored in computational language processing. She went further – this was influence from Wittgenstein – and held that they were infinitely or indefinitely ambiguous, and that only criteria based on meaning could hope to reduce such usage to any underlying machine-usable notation. This emphasis set her off not only from syntactic parsing methods but also from any approach to meaning representation based on formal logic, including any claim to deal with meaning by the use of set-theoretic constructs which never took any serious account of the ambiguity of symbols.

Historically, MMB was vindicated by the growth of semantic parsing techniques during her lifetime and, although syntactic methods have recently recovered the initiative again, one can be pretty sure the pendulum will continue to swing now it is in motion. In recent years, since the work of Montague, there has been an enormous revival of formal philosophical semantics for natural language, in the sense of set and model-theoretic methods, that ignore exactly those ambiguity aspects of language that MMB thought so important. Indeed for many theorists 'semantics' has come to mean just that work, a development MMB abhorred, not because she did not want a philosophical basis for work on language – on the contrary – but because she did not want that particular one.

Formal semantics approaches have not yet proved computationally popular or tractable, and the jury is still undecided on that struggle. It is worth adding that for other languages, particularly Japanese, MT researchers have continued to use semantic parsing methods, arguing

strongly that such methods are essential for an 'implicit' language such as Japanese where so much meaning and interpretation must be added by the reader and is not directly cued by surface items.

BREATH GROUPS, REPETITION AND RHETORIC

These were the three related notions that preoccupied MMB for much of her last 20 years, but which have not in my view yet proved successful or productive, and certainly not to MT where she long sought to apply them. This line of work began when she met Guberina, the Yugoslav therapist who managed to reclaim profoundly deaf persons. From him, MMB developed a notion she later called the Guberina Hypothesis (Masterman, 1963), to the effect that there were strong rhythms underlying language production and understanding (that could be grasped even by the very deaf), and that these gave a clue to language structure itself. From this she developed the notion of a 'breath group', corresponding to the chunk of language produced in a single breath, and the idea that there was, therefore, a phrasing or punctuation in spoken language which left vital structural traces in written language, and which could therefore, be used to parse it by computer. Much time was spent in her later years designing schemes by which the partitions corresponding to idealised spoken language could be reinserted into written text.

From there MMB added the notion that language, spoken and written, was fundamentally more repetitive than was normally realised, and that the points at which the repetition could be noted or cued was at the junctions of breath groups. This notion was linked later to the figures of traditional Greek rhetoric, in which highly repetitive forms do indeed occur, and with the claim that the forms of repetition in text could be classified by traditional rhetorical names.

MMB produced an extensive repertoire of language forms, partitioned by breath groups, with their repetitions marked: a simple standard example would be 'John milked the cows and Mary the goats' which divided into two breath groups, at the beginnings and ends of which were items of related semantic type (John/Mary, cows/goats). Traditional forms of language such as hymns, biblical passages and the longer narrative poems were a rich source of examples for her.

The problem with all this was that it required the belief that all text was fundamentally of a ritual, incantatory nature, if only one could see it – and most people could not. The breath group notion rested on no empirical research on breath or breathing, but rather on the implicit observation that language as we know it is the product of creatures that have to breathe, which fact has consequences even for written text. This last is true and widely accepted, but little that is empirical follows from it.

Almost all linguists agree that spoken language is, in every way, prior to

written. Again, some agree that the phrase is an underrated unit, and language analysis programs have certainly been built that incorporate a view of language as a loose linear stringing together of phrases, as opposed to deep recursive structures. Some support for that view can be drawn from the classic psychological work that shows that sounds heard while listening to text seem to migrate to phrase boundaries. But none of this adds up to any view that language processing requires, or rests on, the insertion of regular metrical partitions carrying semantic import.

Again, the claims about repetition and rhetoric can be seen as an extension of a more general, and certainly true, claim that language is highly redundant, and that the redundancy of word use allows the ambiguity of word sense meaning to be reduced. Programs have certainly been written to resolve semantic ambiguity by matching structured patterns against phrase-like groups in surface text: my own early work did that, and it owed much to MMB's work on semantic message detection. However, the partitions within which such patterns were matched were found by much more mundane processes, including keywords, punctuation and the ends of syntactically-detected phrases, e.g. a noun phrase ending.

The oddest feature of MMB's breath-group work, stretching as it did over many years was that it referred constantly to breathing, but nothing ever rested on that: partitions were always inserted into text intuitively in a way that, to me at least, corresponded more naturally to the criteria just listed (keywords, punctuation, etc.). Finally, of course, it would be overbold to assert that there will never be applications of Greek rhetorical figures to the computer understanding of natural language, but none have as yet emerged, except their explicit and obvious use as forms of expression.

METAPHOR AS NORMAL USAGE

The claim that metaphor is central to the process of language use is one now widely granted in natural language processing and artificial intelligence, even if there are few systems that know how to deal with the fact computationally, once it is granted. MMB always maintained that position (Masterman, 1961, 1980), and the recent rise of 'metaphor' as an acceptable study within language processing is some tribute to the tenacity with which she held it. For her it followed naturally from the 'infinite extensibility' of language use, the majority of which extensions would, at first at least, be metaphorical in nature. It was one of her constant complaints that Chomsky had appropriated the phrase 'creativity' by which he meant humans' ability to produce new word strings unused before, while paying no attention to, indeed positively

deterring study, of aspects of language she considered universal and genuinely creative.

MMB would also welcome anecdotal evidence, of the sort to be found in the work of Cassirer, that metaphorical uses of language were in some historical sense original, and not a later luxury. She rejected the view that language originally consisted of simple, unambiguous, Augustinian names of objects – the view parodied by Wittgenstein (1928, 1972) in the opening of *Philosophical investigations* – but preferred the idea of original primitive atoms of wide, vague, unspecific meaning, which were then both refined to specific referents in use and constantly extended by metaphor. Here, for MMB, was the root not only of the metaphor, but also of metaphysics itself, which consisted for her, as for Wittgenstein, of words used outside their hitherto normal realm of application. But, whereas he thought that words were 'on holiday' when so used, for her it was part of their everyday work.

THE OVERARCHING GOAL: A WITTGENSTEINIAN COMPUTATIONAL LINGUISTICS

There is no doubt that MMB wanted all her theories of language to lead to some such goal, one that sought the special nature of the coherence that holds language use together, a coherence not captured as yet by conventional logic or linguistics. Such a goal would also be one that drew natural language and metaphysics together in a way undreamed of by linguistic philosophers, and one in which the solution to problems of language would have profound consequences for the understanding of the world and mind itself. And in that last, of course, she differed profoundly from Wittgenstein himself, who believed that that consequence could only be the insight that there were no solutions to such problems, even in principle.

It is also a goal that some would consider self-contradictory, in that any formalism that was proposed to cover the infinite extensibility of natural language would, almost by definition, be inadequate by Wittgenstein's own criteria, in just the way that MMB considered Chomsky's theories inadequate and his notion of generativity and creativity a trivial parody.

The solution for her lay in a theory that in some way allowed for extensibility, and in some way justified *ab initio* the creation of primitives. This is a paradox, of course, and no one can see how to break out of it at the moment: if initially there are humans with no language at all, not even a primitive or reduction language, then how can language when it emerges be represented (in the mind or anywhere else) other than by itself? It was this that drove Fodor (1975) to the highly implausible, but logically impeccable, claim that there is a language of thought predating real languages which contains not primitives but concepts as fully formed

as 'telephone'. This is, of course, the joke of a very clever man, but it is unclear what the alternatives can be; or more specifically what an evolutionary computational theory of language can be.

It is this very issue that the current wave of theories labelled 'connectionist' (e.g. Sejnowski and Rosenberg, 1986) seeks to tackle: how underlying classifiers can emerge spontaneously from data by using no more than association and classification algorithms. MMB would have sympathised with its anti-logicism, but would have found its statistical basis only thin mathematics, and would have not been sympathetic to its anti-symbolic disposition.

It is easier to set down what insights MMB would have wanted to see captured within a Wittgensteinian theory of linguistics than to show what such a theory is in terms of structures and principles. It would include that same ambiguous attitude that Wittgenstein himself had towards language and its relation to logic: that logic is magnificent, but no guide to language. If anything, the reverse is the case, and logic and reasoning itself can only be understood as a product of language-users: language is always primary. It is not clear to me whether MMB extended that line of argument to mathematics: I think that she had an exaggerated respect for it, one not based on any close acquaintance, and which for her exempted it from that sort of observation, so that she was able to retain her belief that a theory of language must be mathematically, though not logically, based.

Her language-centredness led her to retain a firm belief in a linguistic level of meaning and representation: she shared with all linguists the belief that language understanding could not be reduced, as some artificial intelligence researchers assume, to the representation of knowledge in general, independent of representational formalisms (a contradiction in terms, of course), and with no special status being accorded to language itself. Indeed, she would have turned the tables on them, as on the logicians, and said that their knowledge representation schemes were based in turn on natural languages, whether they knew it or not.

On the current concern with a unified Cognitive Science, I think her attitude would have been quite different from those who tend to seek the basis of it all in psychology or, ultimately, in brain research. Chomskyans have tended to put their money on the latter, perhaps because the final results (and hence the possible refutations of merely linguistic theories) look so far off.

MMB had little time for psychology, considering it largely a restatement of the obvious, and would, I think, have argued for a metaphysically rather than psychologically-orientated Cognitive Science. Language and Metaphysics were, for her, closely intertwined and only they, together, tell us about the nature of mind, reasoning and, ultimately, the world. She would have liked Longuet-Higgins' remark

that artificial intelligence is the continuation of Metaphysics by other means.

REFERENCES

Amsler, R.A. The structure of the Merriam-Webster Pocket Dictionary. PhD dissertation, Technical Report TR-164, University of Texas, Austin, USA, 1980.

Bar-Hillel, Y. The present state of research on mechanical translation. *American Documentation 2,* 1953, 229-236.

Barwise, J. and Perry, J. *Situations and attitudes.* Cambridge, MA: MIT Press, 1983.

Brouwer, L.E.J. Historical background, principles and methods of intuitionism *South African Journal of Science,* Oct-Nov, 1952, 139-146.

Burton, R. Semantic grammar: an engineering technique for constructing natural language understanding systems. Bolt, Beranek and Newman Technical Report 3453. Cambridge, MA, 1978.

Carbonell, J.G. Metaphor: an inescapable phenomenon in natural language comprehension. Research Report CMU-CS-81-115, Department of Computer Science, Carnegie-Mellon University; also in Lehnert, W.G. and Ringle, M.H. (eds) *Strategies for natural language processing.* Hillsdale, NY: Lawrence Erlbaum Associates, 1982, 415-433.

Chomsky, N. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press, 1965.

Crystal, D. (ed.) *Dictionary of Linguistics and Phonetics,* 2nd ed. Basil Blackwell, 1985.

Fillmore, C. The case for case. In: Bach, E. and Harms, R. (eds) *Universals in linguistic theory.* New York: Holt, Rinehart and Winston, 1968.

Fodor, J. *The language of thought.* New York: Thomas Crowell, 1975.

Frege, G. On sense and reference. In: Geach, P. (translator) *Translations from the philosophical writings of Gottlob Frege.* Oxford: Basil Blackwell, 1960, 56-78.

Katz, Jerrold J. *Semantic theory.* New York: Harper and Row, 1972.

Masterman, M. *Words.* Proceedings of the Aristotelian Society, 1954.

Masterman, M. *The potentialities of a mechanical thesaurus.* Cambridge Language Research Unit, Memorandum, 1956.

Masterman, M. et al. *Agricola in curvo terram dimovit aratro.* Cambridge Language Research Unit, Memorandum ML 84, 1957b.

Masterman, M. *Classification, concept formation and language.* Presented at the Fourth Annual Conference of the British Society Philosophy of Science, 1959a.

Masterman, M. *What is a thesaurus?* Cambridge Language Research Unit, Memorandum ML 90, 1959b.

Masterman, M. Translation. *Proceedings of the Aristotelian Society,* 1961.

Masterman, M. Commentary on the Guberina hypothesis. *Methodos,* XV, 1963.

Masterman, M. Mechanical pidgin translation. In Booth, A. (ed.) *Machine translation.* Amsterdam: North Holland, 1967.

Masterman, M. *Semantic algorithms.* Cambridge Language Research Unit, Memorandum, 1968.

Masterman, M., Braithwaite and Kuhn Analogy-clusters within and without

hypothetico-deductive systems in science. In Mellor, D. H. (ed.) *Science, belief and behaviour*. Cambridge: Cambridge University Press, 1980.
Mellish, C. Implementing systemic classification by unification. *Computational linguistics*, Winter, 1988, 40-51.
Richards, I. A. and Gibson, C. M. *English through pictures*. New York: Pocket books, 1952.
Sejnowski, T. and Rosenberg, C. NETtalk: a parallel network that learns to read aloud. *Johns Hopkins University Electrical Engineering and Computer Science Technical Report*, JHU/EEC-86/01, 1986.
Schank, R. C. *Conceptual information processing.* Amsterdam: North Holland, 1975.
Sparck Jones, K. Synonymy and semantic classification. PhD thesis, University of Cambridge, 1964. Edinburgh: Edinburgh University Press, 1975.
Wilks, Y. Preference Semantics. In Keenan, E. (ed.) *The formal semantics of natural language*. Cambridge: Cambridge University Press, 1975.
Wittgenstein, L. *Philosophische Untersuchungen*. For English translation see Anscombe, G. E. M. *Philosophical investigations*, Oxford: Basil Blackwell, 1972.
Wittgenstein, L. *Tractatus-logico-philosophicus*, London: Routledge and Kegan Paul, 1922.

AUTHOR

Yorick Wilks, Rio Grande Research Corridor, Computing Research Laboratory, New Mexico State University, Box 30001, Las Cruces, New Mexico 88003.