ELSEVIER

Artificial Intelligence 130 (2001) 223–229

Artificial Intelligence

www.elsevier.com/locate/artint

Response

The Cognitive Sciences: A comment on 6 reviews of *The MIT Encyclopedia of the Cognitive Sciences*

R.A. Wilson

Department of Philosophy, 4-115 Humanities Centre, University of Alberta, Edmonton, Alberta T6G 2E5, Canada

As the pluralization in the title of MITECS suggests, and as many reviewers have noted, the stance that we adopted as general editors for this project was ecumenical. We were particularly concerned to generate a volume whose range of topics and perspectives indicated that "cognitive science" was different things to different groups of researchers, and that many even fundamental questions remain open after at least four decades of various interdisciplinary ventures. Implicit in this view is a wariness of any putative magic key to understanding the complexities of cognition in all of its diversity, and the hope that by providing a forum in which this range of work could be reviewed by anyone with time and inclination, the field as a whole would be better positioned to reflect on its future directions.

Readers of the preceding reviews might be interested in a few words about the development of the project. Contracted in the early summer of 1995, MITECS began as a volume projected at half of its eventual size, but with roughly the same scope it has in published form. The general editors, Frank Keil and myself, had been thinking about a volume of this sort independently over the preceding year or so, and so much of the structure of the volume was already outlined by mid-1995. Thus, we were able to move relatively quickly in the second half of 1995 to assemble a team of 9 advisory editors for the six sections that constitute the organization of the volume; as reviewers have noted, the presence of these sections in the print version is manifest primarily by the six overview essays that occupy the first 100 or so pages in MITECS. (I ended up serving as the advisory editor for Philosophy from almost halfway through the project due to the inability of the initial advisory editor to continue and the difficulty of enticing any other sane person to take on the job at that stage.) Articles were commissioned in two rounds, an initial round of around 120 articles late in the Spring of 1996, followed by another round of invitations about a year later, once we had had the opportunity to read over those already written.

E-mail address: rob.wilson@ualberta.ca; rwilson@uiuc.edu (R.A. Wilson).

0004-3702/01/$ – see front matter
PII: S0004-3702(01)00102-3

Pre-production work for the project was completed late in the Spring of 1998, following an extensive period of review and revision. The high level of cooperation we received from authors, as well as the hard work of the advisory editors (and our willingness to forego some articles that were just too long in coming), were crucial to the timely completion of the project.

Links to other MITECS articles within the text of any given article were made largely by authors; links to MITECS articles that appear at the end of each article were added largely by me. After publication, MIT Press staff added to the on-line version the useful links at the end of each article that refer to non-MITECS sources.

The idea of having the majority of articles coming in at around 1000 words or so of text (i.e., about a page) was crucial to containing the project to a single volume while trying to cover all of the ground we deemed relevant to it. Although we added approximately 20% more articles than we initially had planned, the chief increase in length came through reference lists, a policy of allowing authors about a 20% leeway in their assigned lengths, and an increase in the number of articles that we initially projected at 1500 and 2000 words in length. Other aims of the somewhat tight length constraints on individual articles were to facilitate accessibility to a broad audience, and to increase the probability that promised articles would be written within the assigned time frame. In general, the depth of a topic and its coverage within the cognitive sciences are represented in MITECS by the number of articles in any given topic cluster, which are often identifiable from the links for each article. (Since, however, another function of the links is to draw attention to contrastive or more distantly related topics, one can't simply read such clusters off from the links themselves.)

In a project such as this, we didn't expect (and didn't aim) to please all of the people all of the time. But what has been consolation so far is to learn not only that a fairly strong version of the existential quantifier would be in place in restating that adage ("nearly all"), but that many students and researchers alike have found MITECS helpful in their day-to-day work. As the reviews by Carr (on articles relating to ethology) and Dorr (on articles relating to natural language processing) indicate the coverage given to particular topics is, in general, accurate, relatively representative of the field, and at times provocative, even if, as Dorr suggests, there are various ways in which the volume could be further improved. Without discussing particular articles, the review by Okamoto implies the same of the 78 articles that fall under the "Linguistics and Language" section of MITECS, though Lakoff expresses a dissenting view.

Okamoto and (especially) Dorr also point to user-friendly (and not so friendly) features of the on-line version of MITECS. Husbands and Peterson also were able to report on the pros and cons of actually *using* MITECS on-line, something that many readers will appreciate, since at least until the paperback version is issued this summer (at an affordable US$ 65), an on-line subscription (at US$ 25 for 6 months) is the only affordable form of access to MITECS for individual users.

There is a common theme—a surprising one to me in reviews for *Artificial Intelligence*—that emerges from each of the lengthier reviews published here, those by Husbands, Peterson, and Lakoff: that despite our ecumenical efforts, there are shortcomings that turn on, in Husbands's phrase, "the over reliance on computation as a framework to understand and build intelligence". Peterson characterizes MITECS as "a book for believers",

suggesting that "the future credibility of cognitive science may depend on its engagement with computationalism as a useful but perhaps limited and even superficial doctrine", while Lakoff sees MITECS as too firmly anchored to what he calls the "formalist nativist paradigm" to convey much of The Truth (known to aficionados as Cognitive Linguistics). Each of these reviews suggests rather different ways in which MITECS could better achieve its ecumenical goal, but rather than comment on that, let me respond in the time and space available to the general criticism.

It seems worth asking three questions about the centrality of computation or a computational framework to MITECS: How central is that framework to the volume as a whole? Does its presence constitute an indefensible bias?; and, What does this imply for the putative ecumenicism of the volume?

Computation is certainly *a* central notion to MITECS and to the cognitive sciences, both historically and as research in them is actually often undertaken. Its centrality to MITECS is manifest not only in the presence of 81 articles gathered under the "Computational Intelligence" section, but also in computational perspectives represented in all of the other sections, including articles on "computational linguistics" (Avarind Joshi), "computational neuroanatomy" (Eric Schwartz), and "modeling neuropsychological deficits" (Martha Farah). There are three things about this coverage worth underscoring in this context.

First, even counting liberally (which is perhaps how a reader of *Artificial Intelligence* might be inclined to count here), only between one-quarter and one-third of all articles in MITECS cover topics that could reasonably be considered computational in their nature.

Second, as the few examples already cited suggest, the perspectives covered within this general group of articles are amazingly diverse. They cover the application of computation to single neurons (Christof Koch), to psycholinguistics (Dennis Norris), to vision (Ellen Hildreth, Dan Huttenlocher), to cultural representations (Naomi Quinn), and to the use of cognitive artifacts (Ed Hutchins). They cover perspectives perhaps closer to that of "good old fashioned AI" ($\sim$ 40 articles), those that adopt a connectionist orientation ($\sim$ 30 articles), those that represent harder-to-classify, typically newer, perspectives (around 20 articles: examples include "lightness perception" [Alan Gilchrist] and "sentence processing" [Paul Gorrell]), as well as those that discuss general computational notions (examples: "algorithm" [Eric Dietrich], "computation" [Brian Cantwell Smith], "logic" [Jon Barwise]).

Third, many of the articles are critical of traditional and/or more circumscribed construals of the notion of computation within cognitive science. For example, articles on foundational aspects of the role of computation in cognition, such as "computation and the brain" (Pat Churchland and Rick Grush), "computational theory of mind" (Steven Horst), and "rules and representations" (Terry Horgan and John Tienson)—all written by researchers whose critical views of standard computationalist views of the mind are well-known, views reflected in these articles—can be found in the Philosophy section. Articles that mark more radical departures from computational work, such as the vast majority of those in the section Culture, Cognition, and Evolution, are also plentiful.

Which brings me to the second question: is whatever concentration there is on computational frameworks a problematic bias, one that needs to be corrected (as both Peterson and Lakoff imply)? Even in light of what I have said above, one might maintain that such a bias exists. First, one might simply think that MITECS over-represents the

status of computation in the field in general, unfairly displacing some other more central notion within the cognitive sciences. This view is intimated by Lakoff (e.g., see point 8 in his review, more of which in a moment), but it reflects either an unfamiliarity with MITECS as a whole, or a skewed take on the cognitive sciences as a whole, or both. Second, one might think that the centrality of computation distorts the view presented of one or more of the contributing disciplines or its subfields. This, I think, rather than the previous point, is Lakoff's real beef: in particular, that MITECS does not adequately represent the tradition of cognitive *linguistics* and "embodied cognitive science". Third, one might think that MITECS should do more by way of pointing the way *beyond* the computational metaphor in thinking about cognition in a systematic way, and so should downplay computational perspectives even further. This is the sort of view, reading between the lines a little, conveyed at the end of the reviews by Husbands and Peterson.

To take this last point first, there are several problems. For one, it is part of the very ecumenicism underlying the project to arouse scepticism about the claim that there is any one, or even small number, of ways forward across the whole of the cognitive sciences. For two, there is an inherent, conservative bias in a project such as MITECS, which aims primarily to reflect what has been and is being done, achieved, claimed, argued, rejected, and to convey some sense of *how* all of this is happening. (Thus it is MITECS, not the rather less catchy MITEPFDCS, *The MIT Encyclopedia of Possible Future Directions for the Cognitive Sciences*.) We were particularly wary of producing a volume of articles on the latest fads and trends concerning cognition. The only particular future-oriented bias consciously built into the structure of MITECS from the outset was the inclusion of the 50 or so articles in the section "Culture, Cognition, and Evolution". This reflected my own view that the virtual disappearance from cognitive science of anthropological research and, more generally, work in the social sciences, was problematic, and that one corrective would be to include a section in MITECS dedicated to the potentially chaotic cacophony of voices in that kneck of the woods. Other forward-looking themes emerged in some of the section introductions: for example, Stuart Russell and Mike Jordan's view that much of the most interesting work in computational intelligence would move beyond the divide between GOFAI and connectionism, and Tom Albright and Helen Neville's only half-joking prediction that "if cognitive neuroscience fulfills its grand promise, later editions of this volume may contain a section on history, into which all of the nonneuro cognitive science discussion will be swept" (lxix). In short, although MITECS does undertake some constructive work directed at the future directions that the cognitive sciences might take, that was not its chief intended focus.

This brings me to Lakoff's view that MITECS fails to do justice to the perspective of cognitive linguistics, being entrapped—as the *MIT* Encyclopedia of the Cognitive Sciences—within the confines of a "formalist nativist view" of language and cognition. Lakoff's insensitivity to the empty category "Press" following "MIT" in the title of the volume is perhaps understandable; it is something that has, in fact, managed to elude other readers. But speaking of the volume as a whole, the ties of the project to MIT, the institution, are more minimal than one might think. With the exception of Michael Jordan (at UC Berkeley for the last few years but at MIT for the preceding decade), none of the advisory editors and neither of the general editors have primary connections with MIT, either presently or in the past. Fewer than 20 contributors are on the faculty at

MIT, half of these being from the Department of Brain and Cognitive Sciences there. (For comparison, there are more contributors from UCSD and the Salk Institute combined than from MIT.) The only other direct MIT influence on the project was through the influence of the Press's then executive editor, Dr. Amy Brand, whose doctorate is from MIT, and who made several helpful additional suggestions for articles and authors in the area of linguistics and language. This was chiefly by way of calling our attention to some blind spots, once we had commissioned almost all articles. Examples here include "meter and poetry" (Samuel Jay Keyser) and the awkwardly titled "parameter-setting approaches to acquisition, creolization, and diachrony" (Michelle DeGraff). Many readers have found the idiosyncratic nature of such articles, and the articles themselves, thought provoking. This sort of fine-tuning happened in all 6 sections of MITECS and, across all sections, I see no MIT or even "East Coast" bias. The claim that the volume in general reflects a "formalist nativist view" of cognition, particularly one shaped around or by people associated with MIT, is indefensible. In fact, Lakoff's own fixation on the dichotomy between "formalist nativist views" and Cognitive Linguistics is not only antithetical to the ecumenicism of MITECS; it also reflects a yawningly dated view of where the cognitive sciences are at, and doesn't seem to me productive.

Still, let us further explore the idea that MITECS manifests a "formalist nativist" bias. The section Computational Intelligence aside, where such a case might more plausibly be made is with respect to the articles in Linguistics and Language, although I disagree with Lakoff over how deep or problematic this orientation is. In articles that cluster around the topics of phonology and syntax, what we might call "formalist nativist views" are predominant. Even here, however, MITECS offers (a) coverage of a range of accounts of grammar—from "minimalism" (Howard Lasnik), to "head-driven phrase structure grammar" (Georgia Green), to "lexical functional grammar (Mary Dalrymple), to categorial grammar (Mark Steedman)—as well as (b) several lengthier articles, such as "generative grammar" (Geoff Pullum) and "typology" (Bernard Comrie), that express strong reservations about views of language that miff Lakoff. In retrospect, the overall coverage of topics on language strikes me as quite broad and representative, especially once we consider topics (e.g., on aspects of language processing, or on language and culture) that fall under the auspices of sections other than Linguistics and Language, together with those selected within that section.

What Lakoff's review points to, as much by way of exemplification as indication, is the fractious nature of linguistics, something drawn to my attention at several points throughout the project. Linguistics was the only contributing discipline over which a series of battles were fought—from unsolicited, strongly worded recommendations from linguists (not necessarily contributors) either for topics or contributors, to the intransigence of some contributors to round out the coverage offered within their articles. Although this inflexibility was manifest only by a small minority of contributors within Linguistics and Language, it comported with a general pattern of passionate and at times ferocious dedication to one or another One True View on some particular topic in linguistics, a passion and ferocity unmatched by contributors from all other disciplines combined. There are interesting sociological (or perhaps psychological) questions about this phenomenon, for I gather from various sources that deeply felt and expressed disagreements about one

or another One True View are a general feature of professional life in linguistics. (And gee, I thought we got worked up about things in Philosophy.)

As for Cognitive Linguistics (CL) itself, readers now have, courtesy of Lakoff, a sample bibliography of recent work within CL to make some judgments of their own. MITECS itself contains over a dozen articles that discuss CL, including "semantics" (Barbara Partee), "semiotics and cognition" (Patrizia Violi), "language and communication" (Susan Duncan), and "metaphor and culture" (Naomi Quinn). As Lakoff points out, the article "cognitive linguistics" (Karen van Hoek) does a fine job of providing an overview to the CL approach, mentioning and thus (importantly) pointing readers to several key concepts and approaches within CL, including that of construction grammar, mental spaces, and cognitive grammar. While we might have commissioned additional articles on these and other CL-ish topics, that is also true of many other paradigms, approaches, and views, and in our editorial view there was nothing to warrant picking out CL as special here. Had CL emerged as being as central to the cognitive sciences as it is portrayed as being by Lakoff, we think that this would have emerged from the articles by the more than 400 contributors to MITECS. We would then have looked, through last minute fine-tuning, to have provided CL with more explicit coverage than it has, as we did in several other cases. But it didn't, and so we didn't.

Lakoff's own characterization of the state of play here deserves some comment. Although he purports to convey "results" of CL under eight headings, few of these could be said to be results in any reasonable sense of the term. (For example, what sort of result is "4. A Neural Theory of Language has been outlined and is under development", or "6. Language acquisition makes fundamental use of *the opulence of the substrate*:"?) In general, as one moves further through both this set of points and Lakoff's bibliography, matters become more puzzling, even misleading. To take the most extreme case, "8. Applications of these [CL] ideas are currently taking central stage in rethinking traditional disciplines", turns out to correlate with "C. Applications" in the bibliography, where one finds 22 books or articles listed (not bad for coverage of 5 distinct fields), only 2 of which are published work not by Lakoff or one of his co-authors. Speaking for Philosophy, my own field, it is either disingenuous or delusional to represent any of the works listed there as having had more than passing interest to the vast majority of philosophers interested in cognitive science. Of the 79 articles in the Philosophy section in MITECS, few (if any) discuss the work of Cognitive Linguists (though Mark Johnson's work on moral imagination is mentioned in the list of further readings for "moral psychology" [Jon Deigh]). Given that more than 70 of the leading figures in Philosophy with an interest in the cognitive sciences are represented here as authors, one more sanguine about one's own influence than Lakoff appears to be might infer that the message of CL has not (yet?) been taken to heart by such philosophers. Whether this is a "formalist nativist" bias readers can judge for themselves.

Lakoff's review is systematically misleading in another way. The descriptive and bibliographical parts to Lakoff's review both begin with claims and work central to the cognitive sciences—well-represented in MITECS—but go on to point to more contentious and peripheral work as they progress. The descriptive part of the review begins by reminding readers of fundamental work on concepts and categorization (apart from MITECS articles under those titles by, respectively, James Hampton, and Douglas Medin

and Cynthia Aguilar, see also "color categorization" [Paul Kay], "language and culture" [Steve Levinson], and "linguistic relativity hypothesis" [John Lucy]). One direction in which this work led, as Lakoff says, is to CL "and an embodied cognitive science". Now, while the idea of an "embodied cognitive science" *has* been taken seriously within recent work in cognitive science (see, for example, "situated/embeddedness" [Brian Cantwell Smith], "situated learning and cognition" [Colleen Seifert], "dynamic approaches to cognition" [Tim van Gelder], "behavior-based robotics" [Maja Mataric], and "individualism" [Rob Wilson]), CL represents only one of a number of ways to express the idea. It remains extremely unclear either how to unify these competing conceptions, or which, if any, provides the most promising way to develop it systematically. Likewise, there are many ways in which language is shaped by factors beyond the "syntax box", to pick up on Lakoff's "2. Language is also shaped by embodiment", but again the CL development of this idea remains just one amongst many alternative frameworks for doing so.

Turning to the bibliographical section of the Lakoff review, readers might try out the MITECS search engine to explore what sort of coverage MITECS gives to topics such as metaphor and categorization (A.1–A.4 on Lakoff's list), or the work listed under "B. The Neural Theory of Language and Structured Connectionism". They won't, however, find much on Fillmore's notion of framing (A.5), Fauconnier's mental spaces (A.6), or Langacker's cognitive grammar (A.7), as I have already said, for much the reason that they won't find much on "PRO-drop", "situational semantics", or "wide computationalism". Again, readers themselves can judge whether these constitute serious omissions.

As a final reality check, readers might find it useful to see what sort of coverage is given to CL and embodied cognition within not only the volumes that Husbands usefully mentions—those by Osherson, Gregory, Gazzaniga, and Arbib—but also within volumes that are perhaps closer direct competitors to MITECS in terms of their focus, such as Bechtel and Graham's *A Companion to Cognitive Science* (Blackwell, 1998), and, shortly, the projected 4-volume *Macmillan Encyclopedia of Cognitive Science*, edited by Lynn Nadel, currently in progress.