Three books on language: pro-Noam, anti-Noam, net-Noam

William J. Idsardi

Department of Linguistics, Program in Neuroscience and Cognitive Science, University of Maryland, 1401 Marie Mount Hall, College Park MD 20742, USA

Abstract


Three books on language, three general intended audiences, three very different points of view (but all heavily influenced by Chomsky), and (mostly) pitched at three different levels of description and explanation, in the sense of [8].

Yang’s book is aimed at the general intelligentsia, say the readers of the New York Times Book Review, and stays mostly at the computational (problem specification) level, giving us a tour through the state of current thinking within Chomsky’s [1] Minimalist program on how languages are acquired. Yang’s high-level goal is to reconcile statistical learning with the Principles and Parameters approach to Universal Grammar (and its descendents; the technical treatment of this material is in [14]). No actual algorithms or neurons are encountered in the book, though like the other two books, the long-term goal is a vertically integrated theory of language and cognition.

Feldman’s book certainly has the most ambitious title, and it’s intended for those New York Times Book Review readers with PhD’s, who probably also scan Nature, Science and PNAS at least occasionally—the secondary literature crowd. It’s also a reaction to Chomsky (he’s the first author quoted, on the first page of the Preface), and the intent (as the title says) is to take us from the quantum physics of molecular bonds on up to linguistic meaning, and beyond. Contra Chomsky, Feldman wants to study language not as a system in itself, but embodied in its use. This leads to the book’s two credos (page 3, italics in the original):

“Thought is structured neural activity”.

“Language is inextricable from thought and experience”.

We start out at the implementational level, with a brief but nice summary of neurons, fly quickly through the algorithmic level (Feldman “lean[s] toward the low end, with each concept being represented as the activity of a focal...
cluster of 10 to 100 neurons connected to other such clusters” (page 112), and spend considerable time sketching out the problem description for language structure and metaphor, drawing particularly on construction grammar [3] and George Lakoff’s oeuvre (e.g. [7]). Unfortunately, you can’t really get from molecules to metaphor, at least not yet, as Feldman acknowledges (page 334, is this the setup for a sequel?).

The Smolensky and Legendre tomes, with 18 contributors, reminds me most of the Parallel Distributed Processing volumes [12] of the mid-1980s (although the OT Research Group doesn’t get co-author status here), and indeed, I believe this should be seen as showcasing the fruits of the labors of the Optimality Theory [10] research program to reconcile PDP models with generative grammar (Chomsky again, that is) and to give Smolensky’s response to the symbolist-connectionist impasse. This is not light reading, rather, it is intended for those toiling away in the fields, readers of Linguistic Inquiry and Trends in Cognitive Science, et cetera—the primary literature crowd. In these books, we are firmly planted at the algorithmic level (Smolensky’s c-level, page 2:519), with the familiar parallel processing networks—which is the guiding light even when we seem to be doing work on the problem specifications for language. No pretty pictures of axons and dendrites here (for those see Feldman page 51), and we have to wait more than 1000 pages before we get to Fodor and Pylyshyn (1988). That isn’t necessarily a real complaint, though, as we never get to Fodor and Pylyshyn (or Marr for that matter) in Yang or Feldman. There’s no real unification with the implementational level in this book, but Smolensky argues that they have unified the computational and algorithmic levels, and specifically that outer products solve the symbol combination problem (though you can bet that Fodor isn’t buying in).

With the brief descriptions out of the way, you will now know how much trouble you need to go to get the books: Yang’s is at the Borders at the mall, the University bookstore might have Feldman if you’re lucky, and you’re going to have to order Smolensky and Legendre from Amazon.com. Now to the hard part: (why) should you spend the time to read these books?

Yang’s big idea is variational learning on parameters and rules, motivated and explained very succinctly in 20 pages (pages 153–173). Parameter settings are made stochastic, and grammars compete to be selected, with nods to Edelman’s [2] neural Darwinism. Of course, we’ve had this before—Yang [14], but this is a nice presentation, suitable for your friends in other fields, or for your undergraduates. In my opinion Yang has the problem statement exactly right—how do we reconcile stochastic behaviors with parametric grammars? How do we incorporate statistics into models of language without selling out the results from the last half-century of syntactic research? Part of this enterprise is still to separate the wheat from the chaff, to make the hard distinctions between competence and performance phenomena, and to decide what is knowledge OF language (i.e. grammar) and what is knowledge ABOUT language (i.e. using general cognition to reason about objects of language). But part of this enterprise is to face the statistical antithesis directly, and to achieve a synthesis between the symbolic and statistical approaches. Has Yang found the right synthesis? Certainly not yet, but I expect this to be the issue for linguistics for the next period of years, and Yang is in the thick of things. This is also good news for computational linguists and computer scientists generally; CL faced this issue ten years ago [6], so there’s advice to dispense (and maybe even some code to share or sell).

We’ve already seen Feldman’s big ideas (structured neural activity and embodied language), and he’s admirably clear about them. I have to admit I haven’t kept up in this area, but the research program reminds me most of Schank’s work (e.g. [13]) updated to include at least some syntax (Construction Grammar) and some more-or-less relevant brain-imaging experiments. The showpiece proof-of-concept is Regier [11], which does make contact with all three levels (though I imagine the biologists aren’t going to be satisfied with Regier’s simplified visual system), but I’m not optimistic about this scaling up smoothly all the way through to metaphor; this feels more like micro-worlds again, and I think a major determinant in Regier’s success is in finding a tractable (perhaps even moderately encapsulated) sub-problem. That said, there are many useful ideas in the book, but I think Feldman is leapfrogging too high much too soon. For instance, if we want to look for human mirror neurons related to language (Chapter 13), I think we’re much more likely to have success in looking for mirror neurons for phonological features like lip-rounding than for metaphorical uses of grasp (get my drift?). The good thing is that Feldman is clear about his goals and very enthusiastic about the enterprise. This book would make excellent reading in an introductory cognitive science course as it would be certain to spark interesting discussions because the presentation of the material is so accessible (if over-simplified).

The Smolensky and Legendre volume is certainly not over-simplified, and it certainly isn’t as clearly presented. The sidebars (called “boxes”) are intrusive and disrupt the flow of the text—they would work better as appendices—but there’s a ton of material covered here. The material falls into two major groups: neural network computation and linguistics. The big idea in networks computations (the reply to Fodor and Pylyshyn) is tensor products and...
holographic reduced representations (see also [9])—this is neural network recursion. Does it work? I have no idea. Is this the algorithmic basis of locality principles in grammar? Wouldn’t that be nice! Obviously, the jury’s still out on this one, but this has real cachet as a potential computational explanation for essential properties of syntax and sentence processing (but I’m waiting for “Doorknobs are not holographs” or maybe “Chomsky on the holodeck”?). On the linguistics side, we have a variety of contributions in theoretical phonology and syntax, language acquisition and sentence processing. As far as I’m concerned, the big idea here is constraint conjunction, which gives a different way of combining constraints (above and beyond constraint ranking) and introduces the possibility of a formal algebraic structure over the constraint set (even if this possibility is dramatically under-exploited in most OT work). However, constraint conjunction also adds considerable power to the formalism [5], and can’t by itself solve the problem of opacity in OT [1,4]. Although opacity is acknowledged (page 1:506, 524) two pages is about all it gets here. The OT response to opacity (cataloged on page 1:524) is a widening mishmash of computationally intractable additions to the basic theory (constraint conjunction, output-output correspondence, sympathy theory, et cetera), which miss the basic unity of analysis for opacity in the discarded derivational process framework (see also McCarthy 2006 for much discussion and a new proposal, candidate chains). Linguists of the formal persuasion need these books; and anyone interested in the unification of Marr’s computational and algorithmic levels in cognitive science needs to consider Smolensky’s proposals very carefully. Right or wrong, we stand to gain new understanding by critically assessing and testing the tensor product and holographic reduced representations on the fundamental problems in natural language syntax and sentence processing (you can bet that center-embedding and tail-recursion will show up soon in this context).

Finally, on the plus side, all three books are relatively free of arguments about the evolution of language, though they all do manage to work in references to genetics somewhere.

References