

HED: Linguistics

SUBHED: A Voyage into Uncertainty

Joan Bresnan '66

Twenty years ago on a flight back from Europe, I picked up a copy of *Time* magazine and read an article about Panamanian singer and musician Ruben Blades (I happened to have one of his records). One of Blade's remarks struck me, and I copied it down: "I will never be a superstar," he said. "My role is to be different, to do what others won't do, and, as a result, my fortunes will always fluctuate. I will always be viewed with suspicion by some, though not by all, because I move against the current."

This was me, I felt. For most of my career I, too, had been going against the current, swimming upstream, being different.

At Reed I majored in philosophy. Philosophy appealed to me as a matter of temperament, probably, or perhaps because of my inability to remember details for long. I had always had a preference for the persistent, the universal, over things that were merely contingent and accidental, the mere facts. Epistemology, the theory of knowledge, fascinated me: living as we do in a flux of partial information and uncertainty, how do we know anything for sure? I took courses in linguistic philosophy, which dissolved big questions about epistemology and metaphysics into precise (but, inevitably, little) questions about the language we use to pose the big questions. At Reed I also discovered symbolic logic—a beautiful crystallization of the mental, a model for understanding language as a symbolic system.

After Reed and a couple of false starts into philosophy, I ended up in grad school at MIT studying formal grammar in the department of linguistics and philosophy. Formal grammar is an application of discrete math to language, most famously associated with Noam Chomsky. To me it was a similarly beautiful crystallization of the structure of language.

At MIT Chomsky was my doctoral advisor, though my mentor was Morris Halle, who ran the department at that time. (It is a blessing to have a wise mentor who is not in your field of specialization. As with having parents who are not themselves academics, the feeling is that any success you have is undiminished by prior expectations.) The goal of MIT linguists was to infer the nature of the mind's capacity for language from the structure of human language, viewed as a purely combinatorial set of formal patterns, as are the formulas of symbolic logic. The task of formal linguists was to reverse-engineer the structure of language to discover the mental code that produced it. A famous argument given by Chomsky was that probability could never be part of this linguistic theory, because while both of the following sentences are equally unlikely, never before uttered by anyone—

Colorless green ideas sleep furiously.

Furiously sleep ideas green colorless.

—English speakers know intuitively that the first is grammatical and the second (merely a string of words in the reverse order of the grammatical sentence) is not. They know this as part of their implicit knowledge of English, their linguistic competence, which is clearly not based on meaning, since the first sentence is nonsensical yet grammatical.

I loved the symbolic systems that could reveal the hidden deep structure of language. I think what also appealed to me about Chomsky's linguistic theory was that, though couched in the terms of science, it was largely a philosophy of language. Like the epistemologies I had studied at Reed, it had its own theory of knowledge—that the structure of language could not be learned inductively from what we hear; it had to be deduced from innate, universal cognitive structures specific to human language. It did not require a sustained and messy relationship to the world of facts and data. Instead, there was a limitless profusion of data in our own minds, coming from our intuitions about sentences that we had never heard before. With the proper training, scientific research could conveniently be done from an armchair using introspection.

Yet very early on in my career as a linguist I had become aware of discrepancies between this approach to language as a formal symbolic system and the findings of psycholinguists. In fact, starting in 1975 I began to think about designing a more psychologically realistic system of formal grammar in collaboration with a computational linguist from the West Coast via Harvard, Ron Kaplan. We were brought together by Morris Halle and the eminent psychologist George Miller as part of a symposium jointly sponsored by MIT and AT&T to assess the past and future impact of telecommunications technology on society, in celebration of the centennial of the invention of the telephone. What did I know about any of this? Absolutely nothing, but my mentor Morris must have had faith in me. That year I had the freedom to roam mentally into questions about the psychological reality of linguistic theory because I had had the good luck to get a Guggenheim Fellowship.

The next year when I joined the faculty at MIT, Kaplan and I began collaborating, and we soon involved an international group of young researchers in psychology and linguistics (which included Steve Pinker as a postdoc from Canada via Harvard, as well as Marilyn Ford from Australia, Jane Grimshaw from England, K.P. Mohanan from India, and Annie Zaenen from Belgium via Harvard, and other friends and student collaborators). We did psycholinguistic experiments with the new system and began a program of continuing research using it to characterize diverse languages spoken in Australia, America, and Africa, as well as many others in Europe and Asia. The flexible new system—lexical-functional grammar (LFG)—was easy to implement computationally, and it proved useful in industrial linguistic applications such as machine translation projects. Even now, it is the theoretical basis for technology that Microsoft is using to develop queries for its own web search engine, having acquired the LFG-based West Coast start-up called PowerSet, whose motto was, “We parse the Web!”

I got tenure and became a full professor at MIT rather early, but while spending a sabbatical year near Stanford in 1982–83 at the Center for Advanced Study in the Behavioral Sciences, I became attracted to the new Center for the Study of Language and Information (CSLI) that was just then starting up at Stanford. I moved to Stanford in 1983.

The symbolic approach to language that I learned at MIT and pursued in much of my career was intellectually seductive, offering the joys and challenges of puzzle solving and code breaking, along with the optional adventure of travel to other worlds. Some of my students and colleagues were superb field linguists who worked with Australian aboriginal peoples on their own territory and brought forth stunning findings that challenged the conventional designs of grammar. Even I, with my Reed background in

linguistic philosophy and my MIT background in philosophical linguistics, was induced to expand from abstraction to try field work. After moving to Stanford, I began collaborating with Sam Mchombo (from Malawi). I was fascinated, not to say transfixed, by what were to me exotic features of sub-Saharan African languages. These included the use of tones (sounds of different relative pitches) to mark higher-level grammatical structures, and an elaborate gender system with eighteen categories. This collaboration led me to study more African languages and eventually to do field research in Tanzania for a summer in the company of linguist Lioba Moshi. All of this collaborative work on African Bantu languages was well received in the Africanist community, and I was given the unexpected honor of presenting plenary addresses at the Annual Conference on African Linguistics in 1990 and 1995.

Immersing myself in thinking about other, to me “exotic,” languages, I began to see that it wasn’t enough to do linguistics from the armchair. Subjective judgments might be unconsciously biased by the theorist’s preconceptions, or they might be too hastily elicited from brief interactions between poorly communicating strangers. To begin to see even the purely formal patterns of a language, one had to learn how it is pronounced, how it is used in discourses and dialogues, how its speakers interact, because all of the channels of linguistic information are systematically bound up and interconnected. Linguistics did after all require a sustained and messy relationship to the world.

But it was several intellectual shocks arising from my own language, English, that caused me to change my research paradigm completely. The first shock was my discovery that universal principles of grammar may be inconsistent and conflict with each other. The expressions of a language are not those that perfectly satisfy a set of true and universal constraints or rules, but are those that may violate some constraints in order to satisfy other more important constraints, optimizing constraint satisfaction. This insight came into linguistics from outside the field, from neural network approaches to cognition. Yet as my former student Jane Grimshaw pointed out, we can see traces of it everywhere, even in corners of English syntax that had seemed exception-ridden.

In response to this discovery, I developed an optimization-based version of LFG, and soon extended it to models using probability and statistics, as I came to understand the pervasiveness and importance of variation in the structure of language. This kind of move was highly unpopular among Chomskians, who always remember “Colorless green ideas sleep furiously.” Yet I reveled in exploring a new and very different line of research. I particularly enjoyed the seeming paradox of variable grammar as a mental representation of language. At one talk I was asked, “If grammars incorporate random variables, how could you ever know what you were going to say?” I began to recognize that statistics and probability—the mathematics of uncertainty—were epistemology in another guise.

These explorations led to a second intellectual shock as well: I began to realize that we theoretical linguists had no privileged way of distinguishing the possible formal patterns of a language from the merely probable. Many of the kinds of sentences reported by theorists to be ungrammatical are actually used quite grammatically in rare contexts. Authentic examples can be found in very large collections of language use, such as the World Wide Web. Strikingly, the rare, marginal, and “incorrect” construction types in large collections of English language usage parallel the rare grammatical phenomena that can be found across languages of the world. Moreover, judgments of ungrammaticality

are often unstable and can be manipulated simply by raising or lowering the probability of the context. Most remarkably, language users have powerful predictive capacities, which can be measured using statistical models of spontaneous language use. From all these discoveries I have come to believe that our implicit knowledge of language has been vastly underestimated by theoretical linguistics of the kind I had practiced.

But what about Chomsky's famous argument against statistics in grammar? Nearly 50 years afterwards, computer scientist Fernando Pereira showed how a simple statistical model trained on newspaper text can in fact distinguish between grammatical though nonsensical sentences like "Colorless green ideas sleep furiously" and ungrammatical reversals of words like "Furiously sleep ideas green colorless," separating their likelihood of being in the language by five orders of magnitude.

These lines of inquiry were exciting to me in good part because they went against what I had learned previously. They came from moving against the current. And in spite of all this going against the current, I was somehow elected president of the Linguistic Society of America and then fellow of the American Academy of Arts and Sciences. Perhaps the tides were changing.

From one point of view, I have reversed course from the intellectual path on which I started. I completely changed my research paradigm, moving against the mainstream, or at least against the strongest currents in my field. Yet from another point of view, I have been consistent in my love of mathematically precise models of language that reveal its hidden structure and the nature of the mental capacity required to use it. In my early days at Reed, the only available models were discrete symbolic systems like Chomsky's. They required great simplification, abstraction, and idealizations of linguistic data. In recent years, however, new models based on the mathematics of uncertainty have become more widely available to linguists. These provide much more powerful ways to cope with the complexity and variability of language, and these, too, reveal hidden structure and even richer mental capacities.

I moved against the current, yes, but I have always liked being different. Reed attracted many interesting and unconventional young people, withholding grades to prevent superficial status competition, offering a diet of original great works, teaching in a small-class discussion format, and encouraging writing, writing, writing in term papers and in the senior thesis. One of my roommates was an athlete, a beautiful, muscular, and very intellectual swimmer. Another friend was an artist who knew every line in every Shakespearean play. Still another friend, unconventionally conventional in the context of Reed, loved to cook, dress stylishly, and sing the blues. Reed provided a protective habitat for us all.