

## AN INTERVIEW ON MINIMALISM

Noam Chomsky, interviewed by Adriana Belletti and Luigi Rizzi  
University of Siena, Nov 8-9, 1999 (rev: March 16, 2000)

The section on evolution of language begins on p. 21. I have included the rest of the interview for those of you who might be interested; just beware that it gets technical in places. [BV]

### *I. The roots of the Minimalist Program.*

**AB & LR:** To start from a personal note, let us take the Pisa Lectures as a point of departure [1]. You have often characterized the approach that emerged from your Pisa seminars, 20 years ago, as a major change of direction in the history of our field. How would you characterize that shift today?

**NC:** Well, I don't think it was clear at once, but in retrospect there was a period, of maybe 20 years preceding that, in which there had been an attempt to come to terms with a kind of a paradox that emerged as soon as the first efforts were made to study the structure of language very seriously, with more or less rigorous rules, an effort to give a precise account for the infinite range of structures of language. The paradox was that in order to give an accurate descriptive account it seemed necessary to have a huge proliferation of rule systems of a great variety, different rules for different grammatical constructions. For instance, relative clauses look different from interrogative clauses and the VP in Hungarian is different from the NP and they are all different from English; so the system exploded in complexity. On the other hand, at the same time, for the first time really, an effort was made to deal with what has come to be called later the logical problem of language acquisition. Plainly, children acquiring this knowledge do not have that much data. In fact you can estimate the amount of data they have quite closely, and it's very limited; still, somehow children are reaching these states of knowledge which have apparently great complexity, and differentiation and diversity ...and that can't be. Each child is capable of acquiring any such state; children are not specially designed for one or the other, so it must be that the basic structure of language is essentially uniform and is coming from inside, not from outside. But in that case it appears to be inconsistent with the observed diversity and proliferation, so there is kind of a contradiction, or at least a tension, a strong tension between the effort to give a descriptively adequate account and to account for the acquisition of the system, what has been called explanatory adequacy.

Already in the nineteen fifties it was clear that there was a problem and there were many efforts to deal with it; the obvious way was to try to show that the diversity of rules is superficial, that you can find very general principles that all rules adhere to, and if you abstract those principles from the rules and attribute them to the genetic endowment of the child then the systems that remain look much simpler. That's the research strategy. That was begun around the nineteen sixties when various conditions on rules were discovered; the idea is that if you can factor the rules into the universal conditions and the residue, then the residue is simpler and the child only has to acquire the residue. That went on for a long time with efforts to reduce the variety and complexity of phrase structure grammars, of transformational grammars and so on in this manner [2]. So for example X-bar theory was an attempt to show that phrase structure systems don't have the variety and complexity they appear to have because there is some general framework that they all fit into and that you only have to change some features of that general system to get the particular ones.

What happened at Pisa is that somehow all this work came together for the first time in the seminars, and a method arose for sort of cutting the Gordian knot completely: namely eliminate rules and eliminate constructions altogether. So you don't have complex rules for complex constructions because there aren't any rules and there aren't any constructions. There is no such thing as the VP in Japanese or the relative clause in Hungarian. Rather there are just extremely general principles like "move anything anywhere" under fixed conditions that were proposed, and then there are options that have to be fixed, parametric choices: so the head of the construction first

or last, null subject or not a null subject, and so on. Within this framework of fixed principles and options to be selected, the rules and the constructions disappear, they become artifacts.

There had been indications that there was something wrong with the whole notion of rule systems and constructions. For example there was a long debate in the early years about constructions like, say, “John is expected to be intelligent”: is it a passive construction like “John was seen”, or is it a raising construction like “John seems to be intelligent”? And it had to be one or the other because everything was a construction, but in fact they seemed to be the same thing. It was the kind of controversy where you know you are talking about the wrong thing because it doesn’t seem to matter what you decide. Well, the right answer is that there aren’t any constructions anyway, no passive, no raising: there is just the option of dislocating something somewhere else under certain conditions, and in certain cases it gives you what is traditionally called the passive and in other cases it gives you a question and so on, but the grammatical constructions are left as artifacts. In a sense they are real; it is not that there are no relative clauses, but they are a kind of taxonomic artifact. They are like “terrestrial mammal” or something like that. “Terrestrial mammal” is a category, but is not a biological category. It’s the interaction of several things and that seems to be what the traditional constructions are like, VP’s, relative clauses, and so on.

The whole history of the subject, for thousands of years, had been a history of rules and constructions, and transformational grammar in the early days, generative grammar, just took that over. So the early generative grammar had a very traditional flair. There is a section on the Passive in German, and another section on the VP in Japanese, and so on: it essentially took over the traditional framework, tried to make it precise, asked new questions and so on. What happened in the Pisa discussions was that the whole framework was turned upside down. So, from that point of view, there is nothing left of the whole traditional approach to the structure of language, other than taxonomic artifacts, and that’s a radical change, and it was a very liberating one. The principles that were suggested were of course wrong, parametric choices were unclear, and so on, but the way of looking at things was totally different from anything that had come before, and it opened the way to an enormous explosion of research in all sorts of areas, typologically very varied. It initiated a period of great excitement in the field. In fact I think it is fair to say that more has been learned about language in the last 20 years than in the preceding 2000 years.

**AB & LR:** At some point, some intuitions emerged from much work within the Principles and Parameters approach that economy considerations could have a larger role than previously assumed, and this ultimately gave rise to the Minimalist Program [3]. What stimulated the emergence of minimalist intuitions? Was this related to the systematic success, within the Principles and Parameters approach and also before, of the research strategy consisting in eliminating redundancies, making the principles progressively more abstract and general, searching for symmetries (for instance in the theoretically driven typology of null elements), etc.?

**NC:** Actually all of these factors were relevant in the emergence of a principles and parameters approach. Note that it is not really a theory, it’s an approach, a framework that accelerated the search for redundancies that should be eliminated and provided a sort of a new platform from which to proceed, with much greater success, in fact. There had already been efforts, of course, to reduce the complexity, eliminate redundancies and so on. This goes back very far, it’s a methodological commitment which anyone tries to do and it accelerated with the principles and parameters (P&P) framework. However, there was also something different, shortly after this system began to crystallize by the early 80s. Even before the real explosion of descriptive and explanatory work it began to become clear that it might be possible to ask new questions that hadn’t been asked before. But not just the straightforward methodological question: can we make our theories better, can we eliminate redundancies, can we show that the principles are more general than we thought, develop more explanatory theories? But also: is it possible that the system of language itself has a kind of an optimal design, so, is language perfect? Back in the early 80s that was the way I started every

course: “let’s ask: could language be perfect?” and then I went on the rest of the semester trying to address the question, but it never worked, the system always became very complicated.

What happened by the early 90s is that somehow it began to work; enough was understood, something had happened, it was possible to ask the question in the first session of a course: could language be perfect? and then get some results which indicated it doesn’t sound as crazy as you might think. Exactly why, I’m not so sure, but in the last 7 or 8 years I think there have been indications that the question can be asked seriously. There is always an intuition behind research, and maybe it’s off in the wrong direction, but my own judgment, for what it’s worth, is that enough has been shown to indicate that it’s probably not absurd and maybe very advisable to seriously ask the question whether language has a kind of an optimal design.

But what does it mean for language to have an optimal design? The question itself was sharpened and various approaches have been taken to it from a number of different points of view.

There was a shift between two related but distinct questions. There is a kind of family similarity between the methodologically-driven effort to improve the theories and the substantively-driven effort to determine whether the object itself has a certain optimal design. For instance, if you try to develop a theory of an automobile that doesn’t work, with terrible design, which breaks down, say the old car you had in Amherst for example: if you wanted to develop a theory of that car you would still try to make the theory as good as possible. I mean, you may have a terrible object, but still want to make the theory as good as possible. So there’s really two separate questions, similar but separate. One is: let’s make our theories as good as we can whatever the object is... a snowflake, your car in Amherst, whatever it may be... And the other question is: is there some sense in which the device is optimal? Is it the best possible solution to some set of conditions that it must satisfy? These are somewhat different questions and there was a shift from the first question, which is always appropriate (let’s construct the best theory), to the second question: does the thing that we are studying have a certain kind of optimal character? That wasn’t clear at the time: most of these things become clear in retrospect. Maybe in doing research you only understand what you were doing LATER... first you do it and later, if you are lucky, you understand what you were trying to do and these questions become sort of clarified through time. Now you have reached a certain level of understanding, 5 years from now you’ll look at these things differently.

**AB & LR:** You have already addressed the next question, which is about the distinction between methodological minimalism and the substantive thesis. But let us go through the point since you might want to add something. The Minimalist Program involves methodological assumptions which are by and large common to the method of post Galilean natural sciences, what is sometimes called the Galilean style; even more generally, some such assumptions are common to human rational inquiry (Occam’s Razor, minimizing apparatus, search for symmetry and elegance, etc.). But on top of that, there seems to be a substantive thesis on the nature of natural languages. What is the substantive thesis? how are methodological and substantive minimalism related?

**NC:** Actually there is a lot to say about each of those topics: so take the phrase “Galilean style”. The phrase was used by nuclear physicist Steven Weinberg, borrowed from Husserl, but not just with regard to the attempt to improve theories. He was referring to the fact that physicists “give a higher degree of reality” to the mathematical models of the universe that they construct than to “the ordinary world of sensation.” [4] What was striking about Galileo, and was considered very offensive at that time, was that he dismissed a lot of data; he was willing to say “Look, if the data refute the theory, the data are probably wrong”. And the data that he threw out were not minor. For example he was defending the Copernican thesis, but he was unable to explain why bodies didn’t fly off the earth; if the earth is rotating why isn’t everything flying off into space? Also, if you look through a Galilean telescope, you don’t really see the four moons of Jupiter, you see some horrible mess and you have to be willing to be rather charitable to agree that you are seeing the four moons. He was subjected to considerable criticism at that time, in a sort of data-oriented period, which

happens to be our period for just about every field except the core natural sciences. We're familiar with the same criticism in linguistics. I remember the first talk I gave at Harvard (just to bring in a personal example), (Morris always remembers this), it was in the mid 1950s, I was a graduate student and I was talking about something related to generative grammar. The main Harvard Professor Joshua Whatmough, a rather pompous character, got up, interrupted after 10 minutes or so: "How would you handle..." and then he mentioned some obscure fact in Latin. I said I didn't know and tried to go on, but we got diverted and that's what we talked about for the rest of the time. You know, that's very typical and that's what science had to face in its early stages and still has to face. But the Galilean style, what Steve Weinberg was referring to, is the recognition that it is the abstract systems that you are constructing that are really the truth; the array of phenomena are some distortion of the truth because of too many factors, all sort of things. And so, it often makes good sense to disregard phenomena and search for principles that really seem to give some deep insight into why some of them are that way, recognizing that there are others that you can't pay attention to. Physicists, for example, even today can't explain in detail how water flows out of the faucet, or the structure of helium, or other things that seem too complicated. Physics is in a situation in which something like 90% of the matter in the Universe is what is called dark matter -- it's called dark because they don't know what it is, they can't find it, but it has to be there or the physical laws don't work. So people happily go on with the assumption that we're somehow missing 90% of the matter in the Universe. That's by now considered normal, but in Galileo's time it was considered outrageous. And the Galilean style referred to that major change in the way of looking at the world: you're trying to understand how it works, not just describe a lot of phenomena, and that's quite a shift.

As for the shift towards concern for intelligibility and improvement in theories, it is in a certain sense post-Newtonian as has been recognized by Newton scholars. Newton essentially showed that the world itself is not intelligible, at least in the sense that early modern science had hoped, and that the best you can do is to construct theories that are intelligible, but that's quite different. So, the world is not going to make sense to common sense intuitions. There's no sense to the fact that you can move your arm and shift the moon, let's say. Unintelligible but true. So, recognizing that the world itself is unintelligible, that our minds and the nature of the world are not that compatible, we go into different stages in science. Stages in which you try to construct best theories, intelligible theories. So that becomes another part of the "Galilean style". These major shifts of perspective define the scientific revolution. They haven't really been taken up in most areas of inquiry, but by now they're a kind of second nature in physics, in chemistry. Even in mathematics, the purest science there is, the "Galilean style" operated, in a striking way. So, for example, Newton and Leibniz discovered calculus, but it didn't work precisely, there were contradictions. The philosopher Berkeley found contradictions: he showed that in one line of a proof of Newton's zero was zero and in another line of the proof zero was something as small as you can imagine but not zero. There's a difference and it's a fallacy of equivocation; you're shifting the meaning of your terms and the proofs don't go through. And there were a lot of mistakes like that found. Actually, British and continental mathematicians took different paths (pretty much, not 100%, but largely). British mathematicians tried to overcome the problems and they couldn't, so it was a sort of a dead end, even though Newton had more or less invented it. Continental mathematicians disregarded the problems and that is where classical analysis came from. Euler, Gauss and so on. They just said "we'll live with the problems and do the mathematics and some day it will be figured out", which is essentially Galileo's attitude towards things flying off the earth. That's pretty much what happened. During the first half of the 19<sup>th</sup> century Gauss, for example, was creating a good part of modern mathematics, but kind of intuitively, without a formalized theory, in fact with approaches that had internal contradictions. There came a point when you just had to answer the questions: you couldn't make further progress unless you did. Take the notion "limit". We have an intuitive notion of limit: you get closer and closer to a point; when you study calculus in school you learn about infinitesimals, things that are arbitrarily small, but it

doesn't mean anything. Nothing is arbitrarily small. There came a point in the history of mathematics when one simply couldn't work any longer with these intuitive, contradictory notions. At that point it was cleaned up, so the modern notion of limit was developed as a topological notion. That clears everything up and now we understand it; but for a long period, in fact right through the classical period, the systems were informal and even contradictory. That's to some extent even true of geometry. It was generally assumed that Euclid formalized geometry but he didn't, not in the modern sense of formalization, there were just too many gaps. And in fact geometry wasn't really formalized until one hundred years ago, by David Hilbert, who provided the first formalization in the modern sense for the huge amount of results that had been produced in the semi-formal geometry. And the same is true right now. Set theory for example is not really formalized for the working mathematician, who uses an intuitive set theory. And what's true of mathematics is going to be true for everything. For theoretical chemists there is now an understanding that there's a quantum-theoretic interpretation of what they are doing, but if you look at the texts, even advanced texts, they use inconsistent models for different purposes because the world is just too complicated.

Well, all of this is part of what you might call the "Galilean style": the dedication to finding understanding, not just coverage. Coverage of phenomena itself is insignificant and in fact the kinds of data that, say, physicists use are extremely exotic. If you took a videotape of things happening out the window, it would be of no interest to physical scientists. They are interested in what happens under the exotic conditions of highly contrived experiments, maybe something not even happening in nature. Like superconductivity which, apparently, isn't even a phenomenon in nature. The recognition that that's the way science ought to go if we want understanding, or the way that any kind of rational inquiry ought to go -- that was quite a big step and it had many parts, like the Galilean move towards discarding recalcitrant phenomena if you're achieving insight by doing so, the post-Newtonian concern for intelligibility of theories rather than of the world, and so on. That's all part of the methodology of science. It's not anything that anyone teaches; there's no course in methodology of physics at M.I.T. In fact, the only field that has methodology courses, to my knowledge, is psychology. If you take a psychology degree you study methodology courses, but if you take a physics degree, a chemistry degree you don't do it. The methodology becomes part of your bones or something like that. In fact, learning the sciences is similar to learning how to become a shoemaker: you work with a master artisan. You sort of get the idea or don't get the idea. If you get the idea you can do it, if you don't get the idea, you're not a good shoemaker. But no one teaches how to do it, nobody would know how to teach how to do it.

OK, all that is on the methodological side. Then there is a totally separate question: what's the nature of the object that we are studying? So, is cell division some horrible mess? Or is it a process that follows very simple physical laws and requires no genetic instructions at all because it's just how the physics works? Do things break up into spheres to satisfy least energy requirements? If that were true, it would be sort of perfect; it's a complicated biological process that's going the way it does because of fundamental physical laws. So, beautiful process. On the other hand, we have the development of some organ, one famous one is the human spine, which is badly engineered as everyone knows from personal experience; it's a sort of a bad job...maybe the best job that could be done under complicated circumstances, but not a good job. In fact now that human technology is developed you find ways of doing things that nature didn't find; conversely, you can't do things that nature did find. For example, something as simple as the use of metals. We use metals all the time; nature doesn't use them for the structure of organisms. And metals are very abundant on the Earth's surface but organisms aren't built out of metals. Metals have very good constructional properties, that's why people use them; but for some reason, evolution couldn't climb that hill. There are other similar cases. A case that really isn't understood and is just beginning to be studied is the fact that the visual or photosensitive systems of all known organisms from plants to mammals access only a certain part of the sun's energy, and in fact the richest part is not used by organisms: infrared light. It's a curious fact, because it would be highly adaptive to be

able to use that energy, and human technology can do it (with infrared detectors), but, again, evolution didn't find that path and it's an interesting question why. There are at the moment only speculations: one speculation is that there just isn't any molecule around that would convert that part of the light spectrum into chemical energy; therefore, evolution couldn't by accident hit on the molecule the way it did for what we call the visible light. Maybe that's the answer. But if that is the case, the eye is in some sense well designed and in other senses badly designed. There are plenty of other things like that. For example the fact that you don't have an eye at the back of your head is poor design, we would be way better off if we had one, so if a saber tooth tiger was coming after you, you could see it.

There are any number of questions of this kind: how well designed is the object? And no matter how well or badly, to answer that question you have to add something: designed for what? How well designed is the object for X? And the best possible answer is: to let "X" be the elementary contingencies of the physical world and let "best design" be just an automatic consequence of physical law, given the elementary contingencies of the physical world, (so, for instance you can't go faster than the speed of light, and things like that).

A quite separate question is: given some organism, or entity, anything you are trying to study -- the solar system, a bee, whatever it may be -- how good a theory can I construct for it? And you try to construct the best theory you can, using the "Galilean-Newtonian style", not being distracted by phenomena that seem to interfere with the explanatory force of a theory, recognizing that the world is not in accord with common sense intuition, and so on.

These are quite different tasks. The first one is asking how well designed the system is, that's the new question in the Minimalist Program. Of course "design" is a metaphor, we know it's not designed, nobody is confused about that. The Minimalist Program becomes a serious program when you can give a meaningful answer to the question: What is the X when you say "well designed for X"? If that can be answered, then we have, at least in principle, a meaningful question. Whether it is premature, whether you can study it, that's a different matter. All of these things began to emerge after the P&P program had essentially cut the Gordian knot by overcoming the tension between the descriptive problem and the acquisition or explanatory problem; you really had the first genuine framework for theory in the history of the field.

The problems didn't arise clearly until the 50s, although the field has been going on for thousands of years. Until the 50s there was no clear expression of the problem; the fact that on the one hand you had the problem of describing languages correctly, on the other hand you had the problem of accounting for how anyone can learn any of them. As far as I am aware, that pair of questions was never counterposed before the 1950s. It became possible to do it then, because of developments in the formal sciences which clarified the notion of generative process and so on. Once the basic questions were formulated, you had the tension, in fact paradox. The Pisa seminars provided the first way of overcoming the paradox and therefore gave an idea of what a genuine theory of language would be like. You must overcome the paradox. Then there is a framework, and a consequence of that is the rise of new questions like the question of substantive optimality rather than only methodological optimality.

## ***II. Perfection and imperfections.***

**AB & LR:** The minimalist program explores the thesis that human language may be a "perfect system", a system optimally designed to meet certain conditions imposed by other cognitive systems that the language faculty interacts with. But what are the leading ideas about what would count as "perfection"? Some clarification is useful here. One can easily imagine criteria of perfection or optimality according to which human language would be far from optimally designed. Consider for instance the ubiquitous presence of ambiguity in natural language, a property which a superengineer would presumably avoid, given certain goals. One could also argue that language, as an abstract computational capacity, is less than optimally adapted to the human performance

system (with memory limitations, and so on), as it can give rise to all sorts of unusable structures (garden paths, center embedding, etc.), as you have often pointed out. Such criteria of optimal design are a priori conceivable and not unreasonable, but clearly they are not what is intended here. So, what kind of criteria of perfection makes the minimalist thesis sustainable?

NC: Let's distinguish two questions. One is: what do we mean by optimality? Few rules is better than more rules, less memory used in computation is better than more memory used etc. There are some, not precise, general ideas about what optimality is. The second question is: what conditions is the system supposed to meet? I think what you're raising has to do with that question and you're absolutely right: there can be various points of view. If you take a standard functionalist point of view, you would ask: "Is the system designed for its use? So, is it going to be well designed for the uses to which people put it?" And the answer there is "apparently not"; so the system does not seem to be all that well designed for use for the kind of reasons you mentioned (ambiguities, garden paths, lots of expressions that are unintelligible, expressions that are perfectly intelligible but not well formed). In some sense the system is not well designed for use, at least not perfectly designed for use, but it has to be designed well enough to get by. That's all that we discover: it's designed well enough to get by. That raises the question: can we find other conditions such that language is well designed, optimal for those conditions? I think we can, from a different perspective. So instead of asking the standard functionalist question, "is it well designed for use?", we ask another question: is it well designed for interaction with the systems that are internal to the mind? It's quite a different question, because maybe the whole architecture of the mind is not well designed for use. Let me see if I can make an analogy: take some other organ of the body, say, the liver. You may discover that the liver is badly designed for life in Italy because people drink too much wine and they get all sorts of diseases of the liver; therefore, the liver wasn't well designed for function. On the other hand, the liver might be beautifully designed for interaction with the circulatory system and the kidney and so on, and those are just different things. From the point of view of selection, natural selection, things must be well designed, at least moderately well designed for use, well designed enough so that organisms can reproduce and so on. But a totally separate question is: forgetting the use to which the object is put, is it well designed from the perspective of internal structure? That's a different kind of question, and actually a new one. The natural approach has always been: Is it well designed for use, understood typically as use for communication? I think that's the wrong question. The use of language for communication might turn out to be a kind of epiphenomenon. I mean, the system developed however it did, we really don't know. And then we can ask: how do people use it? It might turn out that it is not optimal for some of the ways in which we want to use it. If you want to make sure that we never misunderstand one another, for that purpose language is not well designed, because you have such properties as ambiguity. If we want to have the property that the things that we usually would like to say come out short and simple, well, it probably doesn't have that property. A lot of the things we would like to say may be very hard to express, maybe even impossible to express. You often find that you can't express simple intentions and feelings that you would like to convey to somebody; a lot of personal interactions collapse because of things like that in ordinary life. So, the system is not well designed in many functional respects. But there's a totally separate question: Is it well designed with regard to the internal systems with which it must interact? That's a different perspective and a new question; and that's the question that the Minimalist Program tries to answer.

The way I would like to think of it now is that the system is essentially inserted into already existing external systems: external to the language faculty, internal to the mind. So there's a sensorimotor system which is there, independently of the language; maybe it is somewhat modified because of the presence of language, but in essence it is there independently of language. The bones of the middle ear don't change because of language. And there is some kind of system of thought (conception, intention and so on) which is sort of sitting there. That includes what were traditionally called "common notions" or "innate ideas". Perhaps also analysis in terms of what is called "folk

psychology”, interpreting people’s actions in terms of belief and desire, recognizing things in the world and how they move, and so on. Well, that’s presumably not entirely dependent on language; probably, nonhuman primates have something like that, and perhaps even the capacity of attributing minds to other organisms, a question currently much debated. The language faculty has to interact with those systems, otherwise it’s not usable at all. So, we may ask: “Is it well designed for the interaction with those systems?”. Then you get a different set of conditions. And in fact the only condition that emerges clearly is that, given that the language is essentially an information system, the information it stores must be accessible to those systems, that’s the only condition. We can ask whether language is well designed to meet the condition of accessibility to the systems in which it is embedded. Is the information it provides “legible” to those systems? It is like asking: is the liver accessible to the other systems with which it interacts? If the liver produced something, not bile, but something else that the rest of the body couldn’t make any use of, it wouldn’t be any good; and that’s a different question than whether the liver is well designed for life in a wine-drinking culture. A very different question.

**AB & LR:** An empirically non-vacuous definition of perfection implies the identification of possible imperfections. Inflectional morphology is often referred to as an apparent imperfection. For instance, invented formal languages have a recursive syntax, capable of computing expressions over an unbounded domain, but nothing resembling natural language morphology. What is the driving intuition here?

Morphology seems to be at the same time an imperfection and a defining property of natural languages. How can these two aspects be reconciled within a minimalist perspective?

**NC:** Morphology is a very striking imperfection, imperfection at least superficially. If you were to design a system, you wouldn’t put it in. It’s not the only one, though; no formal language, for example, has a phonology or a pragmatics and things like dislocation in the sense we all understand: expressions appear not where you interpret them but somewhere else, where they might have appeared but don’t in a particular sentence. All of these are imperfections, in fact even the fact that there is more than one language is a kind of imperfection. Why should that be? All of these are at least *prima facie* imperfections, you would not put them into a system if you were trying to make it work simply. A good guiding intuition about imperfection is to compare natural languages with invented “languages”, invented symbolic systems. When you see differences, you have a suspicion that you are looking at something that is a *prima facie* imperfection. There are differences at about every point. Formal languages, for example, don’t have a designated syntax; they just have a set of well formed expressions; the syntax can be anything you like. So, there’s no right answer to the question: What are the true rules of formation for well-formed formulas of arithmetic? “What are the axioms of arithmetic?” The answer is: “Any set of axioms you like to generate all the theorems”. It’s the theorems that are real, not the axioms; the axioms are just a way of describing them, one of many ways. Similarly, if you invent a computer language, it doesn’t really matter which rules you pick to characterize its expressions; it’s the expressions that are the language, not the specific computational system that characterizes them. That’s not the way natural language works. In natural language there is something in the head, which IS the computational system. The generative system is something real, as real as the liver; the utterances generated are like an epiphenomenon. This is the opposite point of view.

Furthermore, the semantics of natural language and of formal languages seem to be totally different, at least in my opinion. Unlike the observation about syntax, which is truism, this thesis is controversial. Not many people agree with me about this, but in my opinion they are totally different. In a Fregean formal system, or in any special purpose system that anyone would construct, the symbols are intended to pick out things, real things. That’s an ideal for natural sciences too. If you construct a scientific theory you want its terms to pick out real things of the world. I mean, if we postulate ECP, we’re assuming there’s something in the world which

corresponds to ECP, that is the purpose of the subject. Scientists may also talk about longitude, let's say, but they know it's not a real thing, it's just a notation for describing things. But it's a goal for science -- and it's built into every invented symbolic system -- that the terms pick out something: that's their semantics, the word-thing relation, essentially. Now, it's a real question whether natural language works like that. I don't think it does. In that case it deviates even in this respect from invented symbolic systems. In fact, it seems it deviates at just about every essential point, and you have to ask why does language have these properties; it is a fair question. A lot of the questions, I think, are too hard... like...if it's true, as I believe, that there's no word-thing relation, the question why there is no word-thing relation is at the moment too hard.

But other questions may not be, like morphology. So let's ask the question why language has morphology, why should language have this apparent imperfection? The primary issue concerns one part of morphology. For example, plurality on nouns is not really an imperfection. You want to distinguish singular from plural, the outside systems want to know about that. So, in fact, plurality on nouns is rather like different words: just as you have "table" and "chair", you have singular and plural, and there are sensible reasons why plural should be an inflection and "chair" shouldn't. Namely, everything has to be singular or plural, but not everything has to be a chair or not a chair. So there are plausible reasons why some part of morphology should be there. Formal languages don't do it but they are just not interested in singularity and plurality, that's not an interesting difference. But human language is interested in this difference, so it has it, like a lexical item, and languages express it as an inflection because of its generality in the system, as distinct from "table" versus "chair", which is not generalizable. So that part is not an imperfection. What is an imperfection is plurality on verbs. Why is it there? You already have it on the noun, so why do you have it on the verb, or on the adjective? Inflection for number looks redundant there, and that is an imperfection. To put it differently, that feature, or that occurrence of the feature, say, plurality on the verb, is not interpreted. You only interpret it on the noun, and that's why in traditional grammars it was always said that the verbs agree with the nouns and that the adjectives agree with the nouns, not conversely. Actually, until very recently from the point of view of generative grammar or structuralist grammar, agreement just looked like a relation. There is no asymmetry to it, no sense in which verbs agree with nouns any more than nouns agree with verbs, one would have thought. And as we know, if you look superficially at languages, it may look as if it is the agreement on the Verb that counts, as in Italian, a Null Subject Language. It looks like it's the inflectional features of the verb that are conveying the information, not of the noun. In fact, there are functionalist studies that reach that conclusion.

If you submit these questions to the minimalist critique, things look quite different. It looks as if there is some real truth to the traditional idea that verbs agree with nouns and not conversely. The thing that is agreeing, presumably the verb, the adjective, the article, and so on, they all seem to have uninterpretable features, features that are not independently interpreted by the outside systems. So, what are they doing there? That's the imperfection. The imperfection is uninterpretable features.

Agreement features are an interesting case, because sometimes they are interpretable and sometimes they are not. But another interesting case is in fact Case. Case systems and inflectional systems have been studied for thousands of years. That's the core of traditional grammar, inflectional systems including Case systems, there's a huge literature on that. By the 1940s and 50s it was getting pretty sophisticated within the structuralist framework. So, say, Roman Jakobson's "Kasuslehre" [5] is a sophisticated interpretation of Case systems. But as far as I can determine, there was never any distinction made between what we now call Structural and Inherent Case; I don't know the literature well enough to check, but I asked other people like Giuseppe Longobardi, and apparently there is no clear recognition of the distinction. In Jakobson's "Kasuslehre", he crucially doesn't make a distinction; his intent is to show that every feature has all the "right" properties (as in the standard structuralist approach), so that each Case feature must have semantic properties. So, Ablative has a semantic property, etc. Then he tries to show that also Nominative

and Accusative have real semantic properties. But, well, they don't. There's a split between the cases that have semantic properties, like, say, Dative, mostly, and the ones that don't, like Nominative and Accusative (or Ergative and Absolutive). As far as I am aware, this split was not noticed until the P&P approach came along; then it suddenly emerged very quickly, in the early 80s, that this core system of natural language, which had been studied for centuries, in fact millennia, broke up into two parts, one of which is an imperfection (at least *prima facie*) and the other which is not. So, the inherent Cases, the ones which are semantically associated, are really not an imperfection: they are marking a semantic relation the interpreter has to know about (like plurality on Nouns). On the other hand, why do we have Nominative and Accusative (or Ergative and Absolutive), what are they doing? They are not interpreted: nouns are interpreted exactly the same way whether they are Nominative or Accusative, and that is like inflectional features on adjectives or verbs: it looks as though they shouldn't be there. This does lead to interesting questions: if you are interested in the Minimalist questions, what you'll ask is exactly that: why are they there? I think there is at least a plausible suggestion: they are there as perhaps an optimal method of implementing something else that must be there, namely dislocation.

The semantics of expressions seems to break up into two parts, at least: what was at one time called Deep and Surface Structure interpretation. It seems there are just different kinds of semantic properties: exactly how they subdivide is not entirely clear, but you can see some differences. There's the kind that have to do with what are often called Thematic Relations, such as Patient, Experiencer, etc.; and there's the kind that look discourse related, such as new/old information, specificity, Topic, things like that. They seem to be different categories of semantic properties, and how to make the break is not very clear. Take quantifier scope; in the work of 25 years ago that was taken to be the prototypical surface property, now it is taken to be the prototypical nonsurface property, LF-property. It's not obvious from the unanalyzed phenomena. But as you learn more, you do see things breaking up into different kinds and then, within the architecture of a more articulated theory, they even seem to appear in different places, assuming the theory is right. So there are the LF-related properties and there are the more surface-related properties. If you look at the surface-related properties, they are typically edge phenomena, they have to do with the edge of the construction. So, say, specificity is typically indicated at the edge of an expression (take Object Shift for instance, a kind of movement to the edge of verbal phrases which yields specificity, old information, etc.). And there is a tradition, which is hard to make clear, but certainly has something to it, which holds that the surface subject tends to be more or less specific; there are exceptions, but it tends to have the specific interpretation. That's perhaps the same point. Real Focus is also an edge phenomenon, in the Left Periphery, and all of these things seem to have in fact some peripheral character. On the other hand, the other category of semantic properties seems to be nondislocated, not at the edge; rather, it involves local relations to other elements that assign the semantic property; a Noun Phrase is related to a verb, a preposition or something like that. That gives the Theta relations. If that's the way the thought system works, there are two kinds of information it is looking for: one edge related, the other locally related. Then, well-designed languages are going to have a dislocation property. An expression will somehow have to distinguish these kinds of information and in fact an optimal way of doing it would just be to resort to dislocation; expressions are phonetically interpreted at the edge even though they are semantically (thematically) interpreted at the local position, the position of Merge. That's a plausible reason, external reason, as to why languages have the dislocation property.

Now, you have to implement the property somehow. How do you implement it? Several things have to be indicated, to make it work. Now we are internal to the computational system. It's as if we had assigned an engineer the problem, "implement the dislocation property", because the system has to do it. So, how do you do it? You have to find the target of dislocation, and it looks as if everything is driven by heads, so let's assume that. If you find a target of dislocation, which will be some head, you have to identify it by some property, which will also determine what kind of element it attracts to it: a Noun Phrase, an interrogative phrase, something else? Furthermore, that

head has to make available a position of dislocation; some do, some don't. And you have to find the thing that is dislocated. So, there must be three things: you need three properties, in technical terms, three features; the term "features" just means properties that enter into the computational system. So, the engineer recognizes: "OK, I need three features": a feature that will identify the target and determine what kind of expression can move to it, one that will identify the thing that is to be dislocated, and one that will decide whether the target has an extra position or not. In fact, the thing that is moved is identified by Structural Case, the target is identified by redundant features -- Agreement features if it is attracting a Noun Phrase -- and the extra position is the EPP feature. What has always been considered weird is the "Extended Projection Principle" (EPP), "extended" because there is no semantic role involved; the role is "here's a position to which you can dislocate", where an element can be interpreted as dislocated. So it seems that you need three features and you have three uninterpretable inflectional features; this suggests, at least, that the uninterpretable features are there precisely to implement dislocation.

There's more evidence for that. One of the properties of the computational system is that, minimally, it has to satisfy the interface condition: expressions have to be interpretable at the interface. You can't have things at the interface that the other systems cannot read. For example, at the sensorimotor level you couldn't have a word that wasn't spelled out phonetically because the sensorimotor system would not know what to do: you couldn't have an orthographic word, for example. And the same is going to be true at the thought end: you have got to eliminate the uninterpretable features. So, somehow the computational system is eliminating all these uninterpretable features, but how will it eliminate them? The natural answer is to eliminate them once they have done their job. If their job is to implement dislocation, then, when they have done it, eliminate them. And it looks as if that is the way in which things work. So, once these features have done their job, they can't do it again: once structural Case has been satisfied, you can't satisfy it again somewhere else. With agreement it's a little more tricky, because there are internal reasons why the system seems to be doing it many times, but once you have taken care of an agreement feature, it can't agree with something higher, for example. It is sort of frozen where it is. All these things hang together in such a way as to lend some plausibility to the idea that these are not imperfections, they are part of an optimal way of satisfying an external requirement, the interface conditions. I don't think this is a knock down argument. It's a plausibility argument but it has some force, and if that is right, then the inflectional morphology turns out to be not an imperfection. Parts of it, like plurality on nouns, are extremely natural, it's good design; other parts like, say, structural Case or agreement features on other elements, they seem to be doing a job that the computational system must carry out and it is a good way of doing it, in fact.

Now, that good way of doing it does lead to oddities: so, for example, sometimes the uninterpretable inflectional morphology functions even though there is no dislocation, with unaccusatives, for example. Suppose we find a structure with a target T that has both (redundant) Agreement features and an EPP feature, but the phrase that agrees with T is unable to move to the target because something else satisfied the EPP feature: perhaps an Expletive, as in (1), or a phrase that is closer to T and therefore preempts the displacement by virtue of locality conditions, as in (2), where  $\bar{t}$  marks the position from which the phrase "to-me" raised to the subject position, satisfying EPP:

- (1) There T-seem (to me) to be many people in the room
- (2) To-me T-seem  $\bar{t}$  to be many people in the room

In English, the rule forming (2) is blocked, but not in other languages; for example Icelandic, or in such Italian constructions as "A Gianni piacciono i dolci," in line with your analysis of experiencer verbs [6]. In such cases, we have "long-distance agreement" of T and the nominal phrase that remains in its initial position, "many people" in examples (1) and (2) (or "i dolci," in the Italian experiencer construction). Visibly, "many people" and "i dolci" agree with the target T (hence

indirectly with the verb that adjoins to T). But according to the account sketched here, the Case -- Nominative Case -- is also assigned as a reflex of this agreement; in some languages, such as Icelandic, the presence of this Case is also visible. In such examples as these, we have all the elements that enter into displacement, but the agreeing nominal is not dislocated. This is a result of blind operation of the mechanisms “designed” to implement displacement, blocked here because other factors intervene.

In case (2) the mechanisms do apply but not to the elements that manifest agreement; rather to the target T and “to-me,” the latter with inherent dative case, expressing a semantic relation that is independent of the Case-Agreement system. Other considerations, still more theory internal, suggest that there is also a kind of “Agreement” between T and the closer raised dative, accounting for the local displacement to satisfy EPP, but only partial agreement, hence not manifested, in accord with general principles.

This is the research direction: try to show that the apparent imperfections in fact have some computational function, some optimal computational function. And there are other cases to be thought about. One massive case is the phonological system: the whole phonological system looks like a huge imperfection, it has every bad property you can think of. Consider the way an item is represented in the lexicon, with no redundancy, including just what is not predictable by rule. So the lexical item does not include the phonetic form in every context, if that is predictable by rule; it just includes what the phonology must know in order to give the output, and it’s a very abstract kind of representation, abstracted from phonetic form. Probably none of the elements that appear in the lexical representation are interpretable at the interface, that is, they are all uninterpretable features. The interface is some kind of very narrow phonetic representation, maybe not even that, maybe a syllabic representation or a prosodic representation. The prosody is not in the lexical item, therefore it is added along the way; what is in the lexical item couldn’t be read at the interface, it has to be modified along the way. Probably the entire phonology is an imperfection. Furthermore the phonological system has, in a way, bad computational properties. For example, one reasonable computational optimality condition is the Inclusiveness Condition, which holds that the computation shouldn’t add anything new; it just takes the features that it has and rearranges them; that is the best system, it doesn’t add junk along the way. The phonology violates it, wildly. The whole narrow phonetics is new, metrics is new, everything is just added along the way. If you look at the phonetics, it seems to violate every reasonable computational principle that you can think of. So, that raises a question: is the phonology just a kind of ugly system? Or is it like what inflectional morphology might be, that is, the optimal solution to some problem? Well, there IS a problem that the phonology has to satisfy, that an engineer designing the language would have to address. There are syntactic structures being generated, and they are being generated the way they are to satisfy the LF conditions, the thought conditions; there is a sensorimotor system, it has its own properties. The syntactic structures have to interact with this “external” system. So, the engineer would be forced to find some way of relating the given syntactic objects to the given sensorimotor system. It would be nice to show that phonology is an optimal way of doing it. That’s a meaningful question, maybe way too hard, but certainly a meaningful question. The best answer that you could hope for is that it is an optimal way of doing it. I suppose that some day it will be possible to turn this into a realistic question, a real research question. A question like this doesn’t even arise until you think of it in these terms, but once it arises it makes a lot of sense, and in fact everything in language can be looked at in this way. The fact that there are parameters ought to follow from something; why didn’t the system just have one state that it could achieve? Why these parameters and not others. There is probably some good reason for that, if we could figure it out.

**AB & LR:** So, the displacement property is an inherent property of natural languages, one that any theory of language aiming at empirical adequacy must express in some way. As for the question why it is so, you offer the speculation that displacement may be an optimal solution to the problem

of having to connect two types of semantic properties to expressions, traditionally, deep and surface semantic properties.

Now, we can pursue the speculation and ask why displacement is the solution chosen by natural language syntax. Clearly there would be other possibilities.

Consider for instance the model, normally adopted in phonology, according to which the sequence of units is on a line at the intersection between distinct planes, such that each plane expresses certain properties, and a unit can be simultaneously assigned properties expressed on distinct planes.

A priori, the integration of thematic and informational properties could work like that, with the same position assigned the property, say, “patient” on one plane and “topic” on another (with, say, deep semantic properties signalled by one kind of affixes, and surface semantic properties also signalled in situ by another kind of affixes). Still, natural language syntax does not seem to work like that in the general case.

Rather, it postulates positions uniquely dedicated to the property “patient” (say, under the Hale-Keyser theory of theta roles), and positions uniquely dedicated to the property “topic”, with the same element occurring in different positions in the same representation, and thus picking up both interpretive properties [7]. This is the displacement property.

In other words, natural languages seem to prefer to solve the problem of connecting deep and surface semantics by proliferating occurrences of elements, rather than by proliferating intersecting planes, or finding other ways to assign different types of interpretive properties to the same position.

Could we speculate on why language systematically goes for this solution? Could this tell us something on the requirements imposed by the interface systems? Could the requirements of linearisation on the PF side be of relevance here? Or some other constraint on the format of legible information on the LF side?

**NC:** It’s a very interesting question, which arises at the outer limits of current understanding, so anything one suggests has to be very tentative.

Suppose first that there was only “deep” semantics, so the problem of displacement does not arise. We now ask: why does language (apparently) identify semantic roles by configuration instead of by particular inflectional elements? Actually, it seems to do both. Thus, Inherent Case (say, Ablative) does identify a semantic role by inflection, while Structural Case (Nominative-Accusative, or Ergative-Absolutive) carries no specific semantic role. For elements with Structural Case, the semantic role is determined configurationally, typically by virtue of their relation to the element that selects them: subject and object of a verb, for example. That this is true is by no means obvious; until quite recently no such distinction was recognized. But it seems to be correct. Furthermore, configurational relations also seem to enter into determining the semantic relation of an element that has Inherent Case.

If so, language uses both devices -- both inflection and configuration -- to assign semantic relations, quite apart from the matter of displacement. We therefore want to know why this is so. The natural place to seek an answer is at the interface between the language faculty and the systems of thought to which it “provides information.” Presumably, these “external” systems distinguish among various kinds of semantic relations, and prefer to have them signalled in different ways. One can proceed to develop further ideas about what these properties of the thought system might be. We are now in a notoriously difficult area, because it is so hard to find out anything about these systems apart from their interaction with the language faculty. We are asking about “thought without language,” in traditional terms, a concept often rejected, though it seems to me reasonably clear that something of the kind must exist.

Turning to the question of displacement, the question about configuration vs. inflection once again arises. Why does language prefer to signal the “surface semantics” configurationally rather than by an inflectional system of the Inherent Case variety? Again, one place to seek the answer is

at the interface. Thus we might ask whether, and if so why, the external systems require that the surface semantics fall together with the deep semantics that is not signalled inflectionally by Inherent Case. But here there are also other possibilities. If surface semantics were signalled by inflection, the underlying morphological system would be complicated. For elements with Inherent Case, there would be double inflection if they have distinct surface-semantic properties; for elements lacking Inherent Case, they would have inflection only in this case. In contrast, if surface properties are signalled configurationally, at the edge, the morphological system is uniform throughout: a single Case inflection always (whether manifested phonetically or not). Possibly that is a factor.

Are requirements of linearisation on the sound side relevant? Perhaps so. To pursue the matter further we should introduce into the discussion languages with more free word order and (typically) richer manifested inflection -- languages of the kind sometimes called “nonconfigurational” (though the term is probably inaccurate).

This is no answer: rather, a suggestion as to where one might look for answers to questions that definitely do arise, and in interesting ways, particularly in the context of serious pursuit of minimalist issues.

**AB & LR:** If it is true that a constitutive characteristic feature of natural languages is to privilege representations with many dedicated positions, each with simple interpretive properties, it becomes important to draw a map as precise and fine-grained as possible of this complex positional system. This is the rationale behind the so called cartographic studies, which are pursued intensely in some research centers in Italy and elsewhere. How can this endeavor relate, in your view, to the topics and goals pursued by the Minimalist Program?

**NC:** This work has led to fascinating results in many areas. To first approximation, the clause seems to be of the general form: [...C...[...T...[...V...]]], where V is the verbal head of the configuration in which deep semantic roles are assigned, T is the locus of tense and event structure, and C (complementizer) is a kind of force indicator distinguishing declarative, interrogative, etc. But the cartographic inquiries have made it very clear that this is only a first approximation: the positions indicated by ... have a rich structure. The “left periphery” includes not only force indicators, themselves differentiated, but also at least fixed positions for topic and focus; and the Cinque hierarchy yields a very detailed and apparently universal array of structures in the T-V region [8]. Other work in progress has provided much insight into the positions at and to the left of T, which host clitics and inflections in various ways; and into apparent parallels between the T-based configuration and the V-based configuration. There are no obvious reasons, at least that I see, why the facts of language should distribute in just this fashion, so once again we are led to the kinds of questions you raised about configurational vs. inflectional solutions, here in a much richer and more diverse terrain.

This kind of work leads us to inquire more closely into the nature of interface relations; the traditional two-interface assumption – sound and meaning – is presumably only an approximation. And beyond that, it leads us investigate the “external” systems themselves, and the conditions they impose on a well-designed language faculty. As is common, these questions have traditional antecedents, but it seems that they can now be addressed on much firmer grounds, and in much more promising ways, in large part as a result of such endeavors as the cartography projects.

**AB & LR:** What kind of empirical discovery would lead to the rejection of the strong minimalist thesis?

**NC:** All the phenomena of language appear to refute it, just as the phenomena of the world appeared to refute the Copernican thesis. The question is whether it is a real refutation. At every stage of every science most phenomena seem to refute it. People talk about Popper’s concept of

falsification as if it were a meaningful proposal to get rid of a theory: the scientist tries to find refuting evidence and if refuting evidence is found then the theory is given up. But nothing works like that. If researchers kept to those conditions, we wouldn't have any theories at all, because every theory, down to basic physics, is refuted by tons of evidence, apparently. So, in this case, what would refute the strong minimalist thesis is anything you look at. The question is, as in all these cases, "Is there some other way of looking at the apparently refuting phenomena?", so as to preserve or preferably enhance explanatory power, where parts of the phenomena fall into place and others turn out to be irrelevant, like most of the phenomena of the world, because they are just the results of the interactions of too many factors. That's one reason why people do experiments. They do experiments to try to get rid of irrelevant phenomena...the point of the experiment is to try to throw out most of the phenomena and discover just those that matter. An experiment is a highly creative act; it's like creating a theory. One may not talk about that in methodology courses, but the working scientist certainly knows it. To try to devise the right experiment is very hard. The first experiment you think of is usually garbage, so you throw out the experiment and try to get a better experiment and so on. Finding the right experiment is very much like finding the right theory and in fact intimately related to it: serious experiment is theory-guided, sometimes to answer questions that arise in the search for explanation and understanding, sometimes because you can see that the phenomena apparently refute your theories and you want to determine whether that is just an artifact. Unanalyzed phenomena don't really matter much in themselves. What matters is the results of properly designed experiments, and "properly designed" means internal to a theory. That's true whether the experiment is about the relation between movement and manifestation of inflectional features, or about language acquisition, or anything else.

Take a concrete example from linguistics and cognitive psychology, one that has been badly misunderstood, the experiment that Bever, Fodor and Garrett did on click displacement [9]. The idea was to see if you could find phrase boundaries, perceptually, by looking at the dislocation of a click. So, you play a piece of tape, put a noise somewhere and ask people where they hear it, and it turns out that they don't hear it where it was, they hear it displaced somewhere; maybe the click was displaced to the edge of the phrase because of some Gestalt property that says that you try to maintain closure, you don't want to be interrupted in a coherent unit, so you perceptually displace it at the edge of the unit. If that worked, it would be an interesting way of finding phrase boundaries. What they were interested in were the hard cases, like Exceptional Case Marking contexts: do you have object raising or not, etc. So if you have "John expected Bill to leave," where is the phrase boundary? Is it after "Bill" or before "Bill"? This is a real question, and the way they proceeded was completely reasonable: first let's design an experiment that works; if we get an experiment which we have faith in, because it is working in the cases where we know what the answer is, then we will apply it in a case where we don't know the answer, and that's what they did. They did a lot of experiments, but what was published was an attempt to show that the experiment works, not to provide new results. In other words, you don't want to have an experiment that is going to give the wrong result in clear cases, i.e. one which in "John saw Bill" would put the break between "saw" and "Bill". First you have to find an experiment that works...Suppose that it turned out that the click invariably got displaced to the middle of the phrase, then it would have been a good experiment, but it would have been interpreted the other way: the Gestalt property is that you displace the click to the middle, we've shown that, because that's what happens. Testing the experiment and deciding how the experiment should be interpreted, that's a large part of the work. In fact, in the case of the click that was essentially all the work. Well, when they got something that seemed to work (displacement to the edge), then they tried it on the hard case: unfortunately, it didn't give very clear results, so it wasn't much pursued. But that shows what experiments are like. Now, this has been seriously misinterpreted. For example, by W. V. Quine, who has been much interested in methodology of linguistics for a long time, since the 1940s. At one time, he argued that phrase boundaries are just an artifact, just as they would be in a formal language, the model he had in mind apparently, as is pretty common [10]. For formal languages, there is no "right" grammar;

it's arbitrary, you pick any one you like. So by analogy, in language the linguist can pick any grammar, depending on one or another concern or interest; the only thing that is real is the utterances. That's a false analogy to start with; human languages are biological objects. What is real – what is in the brain – is a particular procedure for characterizing information about sound, meaning and structural organization of linguistic expressions. The choice of a theoretical account is no more arbitrary than in the case of the visual or immune systems. But pursuing the analogy to formal systems, back around 1970 Quine argued in an article on the methodology of linguistics that it is “folly” to assume that there is a real answer to the question “where the phrase boundary is” in something of the form ABC: it could be between B and C or between A and B. It's just like picking an axiom system for arithmetic, any way you like. Later, after the click experiments came out, Quine changed his mind and said: “Now it's real, because the click experiments show that there really is an answer”. This is a serious misinterpretation. The work on clicks he refers to was testing the experiment, not the phrase structure. If the click experiments had given the wrong phrase structure in clear cases, that would have shown that the experiment is not well designed. One wouldn't say: “The phrase boundaries are not where the linguists thought, they're in the middle of a word” on that basis. Suppose that the click was always heard in the middle of the sentence, so usually in the middle of the word. From Quine's point of view, you'd say: “OK, that's where the phrase boundary is”, but from any scientist's point of view, you would rather say: “Well, it's a terrible experiment”. And in fact, if the clicks were displaced towards the middle of the phrase you just would reinterpret the experiment. From within the framework of the empirical sciences, first you have to test the experiment and that's hard: most experiments are just irrelevant, and to find an experimental procedure that really makes sense is very difficult. It's a theory-internal task, often undertaken because the phenomena of the world are apparently refuting everything, and you want to discover whether, and how, the appearance is misleading.

So, to get back to your question after a long detour. If you want to know what seems to refute the strong minimalist thesis, the answer is just about everything you can think of or pick at random from a corpus of material. That is not particularly interesting because it is the normal situation in the sciences, even the most advanced. Again, this is one of the reasons why people do experiments, which are a crucial part of the “Galilean style”: it is the experiments that matter, and the well designed ones, the ones that fit into a sensible theory. They are the ones that give the data that count, not what you come across. That's not the way linguistics was done until pretty recently. When I was a student, the general idea was to acquire a corpus and try to organize it, to provide a structural description of it. The corpus could be marginally modified by field method procedures -- “elicitation techniques” designed, basically, to determine the scope of partial regularities in observed patterns. But there are no techniques to try to discover data that might be relevant to answering theory-determined queries about the nature of language. That's a creative act. Now, the point of view is that the corpus doesn't matter, it's like the phenomena that you see out of the window. If you can find something in the corpus that is interesting, great. Then you'll explore that with what amounts to doing experiments. But in fact, a lot of the most interesting work has been on things that nobody ever says, like parasitic gaps, for example. You can listen for thousands of years and never hear a parasitic gap, but that's what seems to matter. Sometimes there are really striking results like the work of Dianne Jonas on the dialects of Faroese [11], where she found dialectal differences that nobody had expected and they showed up mostly in things people almost never say, like Transitive Expletive Constructions, and about which speakers are pretty unsure when they say them; but it turned out that there were systematic differences in a category of constructions in areas that people have very little information about, and moreover they weren't aware of such dialectal differences. It's kind of like the parasitic gaps case... Which is, incidentally, normal in experimental sciences: the phenomena that turn out to be interesting are not the normal phenomena of the world, they are usually very exotic.

### **III. Explanatory adequacy and explanation in linguistics.**

**AB & LR:** In the characterization of the aims of scientific linguistics, one important conceptual distinction introduced in the early sixties was the distinction between two levels of empirical adequacy: descriptive adequacy, achieved when a fragment of grammar correctly describes an aspect of the speaker's competence, and explanatory adequacy, achieved when a descriptively adequate analysis is completed by a plausible hypothesis on its acquisition.

The Minimalist Program characterizes a notion of minimalist explanation according to which, to quote from "Minimalist Inquiries" [12], "...a system that satisfies a very narrow subset of empirical conditions in an optimal way -- those it must satisfy to be usable at all -- turns out to satisfy all empirical conditions" (p. 9). Clearly, minimalist explanation is a different concept from explanatory adequacy: explanatory adequacy, in the technical sense mentioned above, could be met by a system not corresponding to minimalist desiderata (for instance, the assumption of an innate list of island constraints could reach explanatory adequacy in certain domains as well as a unifying, simple locality principle, but only the latter would probably meet minimalist standards). How do you see the relations between the two concepts of explanatory adequacy and minimalist explanation?

**NC:** The "list of Islands" model was, of course, developed in some of the most important work of the 1960s. When the tension between descriptive and explanatory adequacy came up, there were several approaches; one approach, which is in "Current Issues in Linguistic Theory" [13] was to try to find principles like A over A, actually also the *wh*- island was in there, and a couple of other things; the other approach was to give a taxonomy of properties, that's basically Ross' dissertation [14], a taxonomy of islands, and an interesting paper by Emmon Bach in which he argued that there should be specific principles for restrictive relative clauses, maybe in all language, and other sets of principles for other constructions. These are just two different intuitions about which way it is going to turn out; and in fact Ross' taxonomy of islands is extremely valuable, a core contribution which everybody goes back to, but that pursues a different intuition, the one that you are describing. What you suggest seems to me quite right. If the truth about language turns out to be something like a system of conditions on rules and constructions, with a unifying locality principle, then only that principle would satisfy minimalist standards, and the program would be a false hope: our explanatory sights simply cannot be set that high -- unless some independent reasons can be found for the other properties postulated, which does not seem very likely -- and core aspects of language would remain unexplained. There also seems to be little prospect for improvement. One would still of course keep to the methodological imperative of seeking the best theory of this biological organ, however "imperfect" it is. My own view is that we can hope for a good deal more than that, but that's a personal judgment.

Assuming so, we might consider a variety of minimalist theses of varying strength. One, which has come up in seminars in Siena, is that every possible language meets minimalist standards. Now, that means that not only the language faculty, but every state it can attain yields an infinite number of interpretable expressions. That essentially amounts to saying that there are no dead ends in language acquisition. You can't set parameters in such a way that you get a system that will fail to have an infinite satisfaction of the interface conditions. That is far from obvious: it is a strong condition on the system. Let's assume that condition is met: minimalist conditions hold for all states of the language faculty, including the initial state. The issue here is not explanatory vs. descriptive adequacy. The standard way to express that distinction is to take a descriptively adequate theory to be a true theory of an attained state, whereas an explanatorily adequate theory is a true theory of the initial state. So, in this view there is a sharp distinction between the initial state, the topic of Universal Grammar, and the attained states, the actual languages. But I think that, at least within the P&P approach, it is more reasonable to forget about that distinction: the language faculty just has states; one state is the initial state; others are the stable states that people reach

somehow, and then there are all kinds of states in between, which are also real states, just other languages. If the strong No Dead End Condition is met, then the minimalist thesis would say that all states have to satisfy the condition of infinite legibility at the interface – and to do so in an optimal manner, to the extent that the strong minimalist thesis holds. That is orthogonal to the dimension of explanatory and descriptive adequacy, because it holds in both the initial state and the attained states. So it's both explanatory and descriptive, but the distinction is kind of put aside. One nice thing about the P&P approach, which at least I didn't realize at that time, is that it essentially eliminates the distinction...it eliminates the principled distinction between the initial state and the attained states. That looked like a principled distinction in the earlier period and it is principled in the sense that the initial state is an expression of the genes, and the others are not entirely, but from the point of view of the adequacy of theories, the distinction doesn't matter...you want an adequate theory for all, they all have to be descriptively adequate, meaning true theories of whatever state you are describing (if it is the initial state, this is what was called explanatory adequacy). If the minimalist thesis holds, it will hold for all states, at least on the "No Dead End" assumption. These questions are really in the process of being formulated, alongside of efforts -- with some success, I think -- to show that strong minimalist conditions can be approached in some domains, sometimes attained.

**AB & LR:** Keeping for a moment this classical distinction, it has often been said that there are tensions between the goals of descriptive and explanatory adequacy as the first typically favors the enrichment of descriptive tools, while the second favors restrictiveness and the impoverishment of the descriptive apparatus. It seems to us that partly analogous tensions could arise between the demands of explanatory adequacy (in the classical sense of "adequacy in addressing the logical problem of language acquisition") and minimalist explanation. It is conceivable that a less structured, hence more minimal, system would allow for more alternative analyses of the primary data, thus making the task harder for the language learner. To give a concrete example, consider a theory of phrase structure permitting a single specifier for each head, and one allowing for multiple specifiers. One could argue, even though the point is not entirely obvious, that the second is more minimal in that it lacks a specification that the first has. But consider the problem from the viewpoint of acquisition: the language learner hears an expression with  $n$  phrases and must integrate them into a structural representation. In the first theory, s/he has no choice: s/he must assume  $n$  heads licensing the phrases as specifiers; in the second theory, s/he has a priori many options ranging from a single head with  $n$  specifiers to  $n$  heads, each with a single specifier. Of course this is crucially related to the question of what can constitute a possible head, and in practice there are many other complications, but the example is simply aimed at suggesting that some tensions could arise here. Do you think this tension actually arises?

**NC:** It could. Minimalist questions are substantive: they ask whether true theories of states of the faculty of language satisfy the interface condition in an optimal way. If a proposal yields as options languages that can't exist, it is just the wrong theory. The same conclusion holds if the proposal does not yield a solution for the logical problem of language acquisition. So, the first condition that has to be met is truth for every state of the language faculty. At the initial state it has been called explanatory adequacy, at a later state, descriptive adequacy. By now, I think this terminology is kind of useless; as I said, just truth matters. Of course, it is not the case that we are given the truth and then we ask minimalist questions: life isn't that simple. You ask minimalist questions to reconstruct your conception of what is probably true and so on and so forth. Logically speaking, the condition in the background must be that you have got the true theory. Take for example the case you mention. There are articles on that in the current literature. *Linguistic Inquiry* has a recent article in which the author says that his way of doing things does not require the special assumption that there are multiple specifiers. But that puts the matter backwards: the assumption that there is a single specifier is a special assumption; to say that there are any number of specifiers is not an

assumption, it's just to say you may continue to merge indefinitely: it merely states that language is a recursive system. To say that there must be a single specifier and no more, is to stipulate that when you Merge twice you have got to start a new category: that's a special enriching assumption. So, there is no issue of getting rid of the extra assumption of multiple specifiers; on the contrary, you would need evidence for the special assumption that you can only have two things attached to a head. Selectional properties of roots may -- in fact surely do -- impose conditions on multiple merge to a single head. But a strong argument would be needed to show that the same condition must be restated, independently, within the theory of phrase structure, complicating that theory, largely redundantly.

In a bare phrase structure theory the distinction between complement and specifier disappears, there is no difference: it's just first Merge, second Merge, third Merge and so on. So, from this viewpoint, a lot of analyses which I have given just don't make any sense. Take adjectives, for instance; I used to worry about whether the element selected by an adjective is a complement of its head or a specifier of its head, sort of different things, but in a bare system you can't ask that question. It's attached to the head, we call it complement if it's first Merge, but it doesn't mean anything, there's no further question to be asked. And the notations that we use are kind of misleading; we put it in front of a head if we mean it to be a specifier, after the head if we mean it to be a complement: those are meaningless distinctions in a bare system. So the whole notion of complement and specifier disappears except as a terminological convenience: you have the things that you merge first, the things that you merge second, and so on.

Let's now assume that we have the simplest system, meaning no extra conditions on how many times you are allowed to merge; so you can do it once, you can do it twice, in which case we call it a specifier, three times, in which case we call it multiple specifiers, and so on, but just merge any number of times you like, plainly the simplest system. And of course we want to know: is it true? Is language perfect in this respect? Or does it have this extra requirement that you can only merge n times for some fixed head, maybe two. Now let's go back to the child acquiring the language. If the child is acquiring the language with the principle of Universal Grammar that says you can merge as many times as you like, the child hears two merges and OK, that's fine; then he hears the third thing come along and, you're right, that the child has two choices. One is to say: "OK, it's third Merge", the other is to postulate a new head. But that's a hard choice: to postulate a new head you have to have evidence, you have to know what head it is, to find it somewhere and if it is a zero head as it could be in this case, it is very hard. If it is a head that doesn't have any semantics, you are in trouble because that head will have to disappear in the course of computation, which will leave you with a headless category and you'll have to tell some story about that. If there is some universal set of options, say, Cinque's hierarchy, then you can pick something out of that, but then there has to be a semantic consequence and you have to have evidence for it. So I don't think it's a question of harder or easier choice, it's just different choices. If the Universal Grammar has Cinque's hierarchy and no limitation on merging, then when you get to that third element, the child will have to ask whether it has the semantics of something in the hierarchy. If it does, then that's where it belongs; if it doesn't, just Merge down below and that's the answer.

Let's now take the other approach; suppose that phrase structure theory is complicated to impose the (largely redundant) requirement of single or double Merge, not triple Merge. Then the child is forced to find another head; and if there is nothing around that makes any sense, it will just have to invent it, and that's a harder task. So, I don't think that the conflict breaks up this way. It seems to me that there are different factual assumptions about the nature of language. Are there heads available with the kind of semantics that will compel the child to merge to them, whether it is third Merge or fourth Merge?

In fact the same question arises for second Merge. Suppose the child assumes first Merge on a head, and then a second expression comes along. Let's assume a Universal Grammar which has no limitation on specifiers and the Cinque hierarchy. After the first Merge, when the second expression comes along, the child is confronted with the same question: does this have the

semantics of one of the positions of the hierarchy, because it has some kind of aspect interpretation, or the like? Well, if so, then the child should postulate a new head; if not, then the element is a specifier of the first head and the same question arises on third Merge, fourth Merge, and so on. The situation you are mentioning could arise and then it would be a question of truth; so the truth may be that you have more complicated phrase structure, with conditions on the number of specifiers over and above those that follow from selectional requirements. For example, take the LCA (Linear Correspondence Axiom [15]). If that theory is true, then the phrase structure is just more complicated. Suppose that you find out that government is really an operative property. Then the theory is more complicated. If ECP really works, well, too bad; language is more like the spine than like a snowflake [16]. You can't change reality, you can only ask: "Does reality happen to meet these surprising conditions?"

#### ***IV. Minimalist questions and other scientific domains***

**AB & LR:** Granting the common background of methodological minimalism as a component of scientific inquiry, are substantive minimalist questions ever asked in other scientific domains?

**NC:** Not often, I suppose, but they are in some. So, for example, there is a sort of a standard joke in physics and mathematics that the only numbers are 1, 2, 3 and infinity; the others are too complicated, so if anything comes out, say, seven, or something like that, it is wrong. And in fact that actually shows up in scientific work, it showed up in the development of the theory of quarks, apparently: if I remember correctly, when Murray Gell-Mann and his associates were devising the theory, it turned out that they had evidence for seven quarks, but nobody was happy with that, because 7 is too ugly a number; so the assumption was that the picture must be reconstructed in terms of 2 and 3, which are nice numbers. And after further experimental work stimulated by that intuition, the prettier picture turned out to be true. I think that that kind of reasoning does go on. In a sense the discovery of Pluto was kind of like that. There were perturbations, so it could be that the world is ugly and you have to make up some story; but everyone was very happy when they found a postulated entity out there which may or may not be a planet, that is debated, but whatever it is, it is out there and it accounts for the perturbations without complicating physical theories. You want the systems to look nice. Take the Periodic Table, for example. The known facts didn't entirely fit, but it was so nice that it had to be right, so it didn't matter if they didn't fit. There are famous examples in the history of science that are kind of like this. Chemistry, which is a rather revealing model for linguistics, provides many examples. Many chemists were unhappy with the proliferation of elements and chemical atoms in the theories of Lavoisier and Dalton. Humphry Davy, for example, refused to believe that God would have designed such an ugly world. At the same time, in the early 19<sup>th</sup> century, William Prout observed that the atomic weights of the elements were pretty close to integral multiples of the atomic weight of Hydrogen, and fudged the data to yield whole numbers exactly. "Prout's hypothesis", as it was called, stimulated heavy experimental inquiry trying to find the exact deviation of the atomic weight of heavier elements from an integral multiple of Hydrogen and to try to find some explanation: is Prout's hypothesis right or wrong? Are all elements constructed from Hydrogen, as he speculated? Finally isotopes were discovered in the 1920s and then it all became clear: it was clear that Prout's hypothesis was fundamentally correct. Without an understanding of isotopes and atomic theory generally, the data are a mess. But if you reanalyze the data in terms of new theoretical understanding, you discover in just what sense Prout's hypothesis was correct, because you get a proton, many protons, its integral multiples, and electrons don't add much, and isotopic effects modify the numbers systematically. The research inquiry was driven by the hope that somehow this pretty law will turn out right and there will be a reason for it; finally the reason was found, and incidentally a good deal of the experimental work of a century went out of the window; nobody cared anymore what the average deviations were because you had a fundamental explanation for them.

I suppose the Galilean ideal of perfection of nature is, at some level, a driving force in all inquiry, but it certainly isn't very much of a leading force in most fields, any more than it has been in linguistics. A good reason is that it is so hard to gain something approaching descriptive adequacy that you can't realistically ask further questions.

Take a look for example at Mark Hauser's recent comprehensive study *Evolution of Communication* [17]. It really is a comparative study of communication, comparing communication systems. He reviews a lot of systems and describes them in very delicate detail. Take the bee dance. There are extremely detailed descriptions of it, but it's basically like descriptive linguistics. Questions that go beyond are apparently too hard: for example, what is the "generative grammar" of the bee dance, the internal state that allows for this range of dances and not some other range? Or questions about neural mechanisms, their role in action and perception, their evolution. The problem of just giving a description is hard enough; and then finding some understanding of the function of the dance. To go beyond that, to get real minimalist questions is hard, but there were people who were trying to do it also in biology. A famous example is D'Arcy Thompson...

**AB & LR:** This leads to the next question. Let us assume that some form of the minimalist thesis is correct, and human language is a kind of optimally designed system. You have often stressed that this would be a very surprising conclusion in the context of biological systems, which are characterized by the "bricolage" or tinkering of evolution, in François Jacob's terms [18]. So, it would be useful to try to spell out the consequences of this discovery for biology.

One possible line of approach could be to think that language is effectively rather unique among biological systems, possibly in relation to its combinatorial character; but it could also be that language readily reveals something that is more common than usually assumed in biological systems, but only difficult to detect. Could it be that the role of tinkering has been overstated? And that at different levels of the evolutionary scale "perfect systems" may have come to existence, but are hard to tease apart from their biological context?

**NC:** That is, I think, quite reasonable. It is unpopular today, but the fact is that if you take a look at anything that you don't understand it's going to look like tinkering. That was true of the way people looked at languages...If you go back to the 1950s a standard assumption – I am paraphrasing Martin Joos, one of the major theoreticians -- was that languages can differ from one another without limit and in arbitrary ways. Basically, there is nothing much to say about language: almost anything goes [19]. That is certainly what it looks like. If you consider the range of languages in the world, it looks as though you can find just about anything. That was a standard point of view in structuralist linguistics, which departed from this assumption only in limited ways: there is some fixed structure of the phonemic system and maybe a little bit more, maybe some of the morphology, some loose conditions on phrases...but essentially anything goes. Sapir said similar things and in fact it's pretty common [20]. And it's true: if you look at anything that you don't understand it is exactly what it is going to look like. With regard to evolution, everybody believes Darwin is basically right, there's no question about that; but beyond that, not too much is understood. For evolution of species, there are few cases in which it can be demonstrated, by the standards of the sciences, that natural selection operated, though everyone assumes that it is true. It is not easy to measure selective advantages of traits. When you look at what are called "natural selection explanations," what you often find is something different. Hauser's book is a good source there. He's trying to show in detail what everybody believes generally: that natural selection functions crucially to yield and design an output. But the kind of argument that he gives doesn't show that. So he takes bats and shows that they have an amazing technique of echolocation: they can find an insect flying somewhere and shoot right at it by some kind of echoes that man-made systems can't duplicate. The conclusion is: look at how beautifully natural selection worked. That is very plausible, but the argument doesn't show it; what is shown is that it has these beautiful characteristics. A recent review of the topic in *Science* points out that it is plausible to suppose that

piranha teeth evolved for cutting, “but we have no direct evidence that that was the case.” A creationist might say, irrationally, that God made it that way. It is just that if you have a naturalistic approach to the organic world, you assume that it must have been largely the result of natural selection. A description of the beautiful adaptation to the organism’s needs is just formulating the problem to be addressed. The problem is: here’s the object, here are its strange properties marvelously adapted for survival and reproduction. That sets the problem, but doesn’t answer it. It is often taken to be an answer to the problem, on the assumption that the outcome has to be the result of natural selection. The dogma in this case is pretty plausible (it’s hard to think of anything else), but that’s not an answer and sometimes, when things have been looked at carefully, the answer turns out to be something different and unexpected. Things are what they are, not necessarily what we dreamt of. In fact, at the moment, little is known about evolutionary processes other than the main principles, and a huge amount of descriptive work that yields highly plausible assumptions (like echolocation and piranha’s teeth), of course, a lot of special things about what genes do, and so on and so forth. But it does look mostly like a mess, and it may not be. It may be that the whole of evolution is shaped by physical processes in a deep sense, yielding many properties that are casually attributed to selection.

Now, of course, when people say that something is the result of natural selection they don’t mean it literally. Natural selection can’t work in a vacuum; it has to work within a range of options, a structured range of options; and those options are given by physical law and historical contingency. The ecological environment is in a certain state and it is going to impose constraints: you could imagine a planet in which you have different ecological conditions and things would work in a different way. So, there are contingencies and there’s physical law and within that range natural selection finds its way, finds a path through it; but it can never be the case that natural selection is acting on its own. The logic is rather like that of behaviorism, as was pointed out by Skinner, incidentally. He thought it was an argument for his radical behaviorism, that it works like unstructured natural selection: so the pigeon carries out any possible behavior and you reinforce the one you want, and you get pigeons playing ping pong, etc. He argued this is the same logic as natural selection, which is true, but what he missed is the fact that natural selection requires a structured environment, structured entities, and the conditions imposed by natural law, and the same is true of the pigeon. So, it is the same logic, and the same mistake for both. And it’s common... When you read these excited pronouncements about “show me good design and I’ll find natural selection,” ...”God or natural selection” -- taken literally, it’s worse than Creationism. Creationism at least is coherent; you can be a rational creationist (Voltaire, Jefferson,...), you can even be a neo-Darwinian. A rational creationist could say OK, all this stuff happened by natural selection but God was necessary to do X. There is no point in this vacuous assertion but it is not incoherent. On the other hand, a belief in pure natural selection would be totally irrational; it is assuming that some selectional process can take place in a vacuum, which can’t happen. It is always the case that what goes on is to some extent conditioned by physical law at least. There is a kind of a “channel” set up by physical law and, in addition to that, there are historical contingencies and so on. Within those structured constraints, natural selection can operate. Well, that raises a question, always: to what extent is the channel functioning in determining the output? It is going to be more than zero, it has to be. In some cases, it may approach 100%. Take the fact that you find the Fibonacci series showing up all over the place. Nobody believes that it is either God or natural selection; everybody assumes that it is the result of physical law and by now there are nontrivial physical explanations of why you should find it. So, between 100% and something, that’s the effect of the “channel”.

Now, when you understand very little and it all looks like a mess, you assume...OK, it is just wandering through the space of possibilities, it is tinkering. But, as you learn more, you may find out that it is not true at all, maybe most of biological evolution is like the Fibonacci series. There is a tradition in modern biology of serious scientists who have tried to exploit that idea. The most famous one is D’Arcy Thompson [21], who tried to show that you could account for large aspects of the nature of organisms by looking at biophysics, basically: what kinds of forms could

there be? Actually Goethe did something similar [22]. He had interesting ideas, some of which turned out to be right, I mean, not the way he thought, but basically right: in plant growth everything is a replication of the same structure over and over again, the stem and the leaf; he kind of guessed, it's a mixed story, but kind of right. With D'Arcy Thompson this becomes real science. Not much was done with it, because it probably was too hard. But that opened a tradition. The next famous person who picked that up was Alan Turing [23]. This is not too well known outside biology. Turing is mostly known for his mathematics, but he also worked on biological problems. He was a serious scientist and he was interested in showing how, if you have a thermodynamic system of some kind and some singularity exists, a slight perturbation, it might lead to a discrete system, suddenly. So, he was interested in things like zebra stripes: how come zebras have stripes instead of just some mess? And he tried to construct models in which you'd get things like zebra stripes, just out of physical processes with a tiny perturbation which changes things around. And the mathematical models are apparently right, so I am told. The question whether it works with zebras is another problem, I think the current belief (I am no expert) is that for zebras it probably doesn't work, but for angelfish it probably does work. There's a certain kind of fish that has some weird stripes all over the place and apparently the Turing models or some modification of them do a reasonable job for explaining that.

At the level of very simple systems, a lot of this is pretty much assumed. Mitosis is a case in point; nobody thinks there are genes that tell the breaking cell to turn into spheres, just as you do not have a gene to tell you to fall if you walk off the roof of a building. That would be crazy, you just fall because physical laws are operating, and it is probably physical laws that are telling the cells to break up into two spheres. Well, another case that is generally assumed is the shell of viruses, which are polyhedrons and in fact icosahedrons. It turns out just by pure geometry that there are only certain kinds of forms that can appear and be stable and fit together. The viruses pick one of those forms and they pick the one of the possible geometrical figures that is closer to a sphere, so they don't pick pyramids, they pick icosahedrons... Maybe that involves selection, but the possible viral shells are assumed to be determined just by physical law. Or take the honeycomb of bees, which is again based on polyhedrons... There are other things: there is an organism -- nobody even knows whether to call it an organism -- called a slime mold, which begins with little organisms and they all hang together and then become a bigger organism, and finally they split up and become separate organisms... This happens in a regular fashion and I understand that the mathematics of this is pretty well worked out. There is some fairly straightforward physical property that will lead to this complicated-looking behaviour once it is operative. So, superficially, that might look like tinkering and fitting some environment, but in fact it is probably just some slight change that led to this happening. How far does that go? Most things are just not understood, so you don't know how far it goes. When you go beyond simple structures, you are guessing what might have happened, and when something is learned, the guess often turns out to be wrong, because you just can't guess, there are too many possibilities, many not yet even imagined. The evolution of the eye, for example, has been extensively studied, and a standard conclusion was that it had evolved independently about 50 times. Recent work has found that there is a single origin, and a single "master control gene" for all eyes in the organic world [24]. Then, over billions of years, evolutionary processes (natural selection functioning within a structured "channel") gave rise to many kinds of eyes, superficially very different, but with deep uniformities.

Now let's turn to language. It appears to be a fact that language is biologically isolated. Let's look again at Hauser, which is THE encyclopedic study of evolution of communication, a study of comparative communication, really. Language doesn't even fit in his taxonomy. Human language is the exciting topic, so the book starts with language, it ends with language, and in between there's comparative communication studies. But in it there is a taxonomy of possible systems and language does not belong. The possible systems include nonhuman primate calls, bird songs, etc.. There are systems related to survival, mating and reproduction and there are those involved in the identification of the caller and so on. That's about it. Language doesn't fit in. You

can use language to identify yourself, for reproduction, for warning about predators. But one can't study language seriously in these terms. Language simply has no place in the taxonomy. In fact, Hauser kind of mentions this, but without making clear the consequences of what he is saying. He says that everything in his book is "irrelevant to the formal study of language"; well, "irrelevant" is too strong, but that is his statement. But what is the formal study of language? The answer is: virtually everything about language. He may have in mind rules in some notation, but it isn't that..."the formal study of language" includes all the work that seeks to determine the nature of language, just as "the formal study of bee dance" includes virtually the entire literature on the topic. So, whether it is syntax or semantics or phonology or pragmatics or whatever you call it, that's the formal study of language. If everything in the book is "irrelevant to the formal study of language," it is just another way of saying that language doesn't belong in this taxonomy. And apparently that's true. He is certainly trying to make a serious effort to show that language belongs, but when you look, it turns out that it doesn't fit, whether we have in mind the properties of language or its various "functions." When Hauser gets to the last chapter of the book, called "Future Directions", he speculates about how we might some day be able to say something about evolution of these systems, because now we can say essentially nothing. In the case of language, what he says is: "look, there are two problems; you obviously have to memorize a lot of words and you have to have a generative system, which is going to give you an infinite array of expressions, so something has to deal with those." Well, how do you do it? The infinite array of expressions, he just drops: he mentions the problem with no speculation, which makes sense, because there is no serious speculation. What about the explosive growth of the array of words? He observes that there is little to say about this, either. It is not like animal calls. Word learning, he points out, must involve a capacity for imitation; so humans have an innate capacity for imitation. Of course, far more than that, as he recognizes. What about the capacity for imitation, then? Well, that turns out to be a total mystery too. According to Hauser, that is not found in any relevant form elsewhere in the organic world, and there's no way of knowing how that came about, so he (in effect) concludes. So, it's a total dead end. There is essentially nothing to say, language is off the chart. That is the basic conclusion that follows from his comprehensive review of comparative communication.

That doesn't mean that language is not the result of biological evolution, of course we all assume it is. But what kind of result of biological evolution? Well, here you have to look at the little bit we know. We can make up a lot of stories. It is quite easy...for example, take language as it is, break it up into fifty different things (syllable, word, putting things together, phrases and so on) and say: "OK, I have the story: there was a mutation that gave syllables, there was another mutation that gave words, another one that gave phrases..." Another that (miraculously) yields the recursive property (actually, all the mutations are left as miracles). OK, maybe, or maybe something totally different; the stories are free and, interestingly, they are for the most part independent of what the language is. So if it turns out that language has a head parameter, same story; if it doesn't have a head parameter, same story. The story you choose is independent of the facts, pretty much. And that's going to be the case until you know something. You can make up stories about the eye, about wings and so on. What happened is what happened, it is not necessarily the story you chose. And looking at the marvelous adaptation of some system to its environment, when that is what we find, just sets the problem, it is not the answer, contrary to common misunderstanding.

Going back to language, what you have is a system that is, as far as we know, essentially uniform. Maybe there was some speciation at one point but only one species survived, namely us; there seems to be no variation in the species. True, we find Williams' syndrome and Special Language Impairment. But that's not variation in the species in any meaningful sense: those are deviations from the fixed system that occur now and then, but the basic system seems to be uniform. In other words, kids learn any language anywhere, as far as we know, which means the basic system is uniform. Nobody has found any genetic differences; maybe there are some, but they are apparently so slight that we can't detect them. So, it is a fundamentally uniform system, which means that since its emergence there has not been any significant evolution. It has just stayed that

way. People have scattered, there are groups of people that have been separated for a long period, but nobody can detect any language difference. So it's apparently a recent thing, too recent to have undergone much evolution. There is also a point that Jerry Fodor has recently stressed [25]: language is different from most other biological systems, including some cognitive systems, in that the physical, external constraints that it has to meet are extremely weak. So, there's some innate system of object recognition: infants can identify object constancies; they know things don't go through barriers, etc. But that system, whatever it is, has to be attuned to the outside world; if you had a system that had objects going through barriers and so on, you couldn't get along in the world. So that system is sort of controlled by the outside world. Then it makes sense to speculate that it was selected...this is a speculation, but plausible, like echolocation. On the other hand, language doesn't have to meet that condition, or it has to meet it in an extremely weak way. You have to be able to talk about the world, somehow, but there's any number of ways of doing that. The fundamental condition that language has to meet is that it can be used, that the person who has it can use it. Actually you can use language even if you are the only person in the universe with language, and in fact it would even have adaptive advantage. If one person suddenly got the language faculty, that person would have great advantages; the person could think, could articulate to itself its thoughts, could plan, could sharpen and develop thinking as we do in inner speech, which has a big effect on our lives. Inner speech is most of speech. Almost all the use of language is to oneself, and it can be useful for all kinds of purposes (it can also be harmful, as we all know): figure out what you are going to do, plan, clarify your thoughts, whatever. So if one organism just happens to gain a language capacity, it might have reproductive advantages, enormous ones. And if it happened to proliferate in a further generation, they all would have it. In a larger group all that is necessary is that it be shared. The connection to the outside world is extremely weak and therefore it could be very stable, because there is no point in changing it; there's no advantage to any change that takes place, or it could be stable because it just didn't have enough time. One way or another, it has evidently been stable.

What happened pre emergence? That's anybody's guess; it seems to be absurd to regard it as an offshoot of nonhuman primate calls. Language doesn't share any interesting properties with them. Or with gestural systems; or anything that we know about; so, you are stuck. Language has highly unusual properties: discrete infinity is unusual, displaced reference is unusual, the most elementary structural and semantic properties seem unusual. It is possible that what happened is what Richard Lewontin and others have speculated [26]: the brain was exploding for a million years; it was getting way bigger than among other surviving primates, and at some stage (for all we know about 100,000 years ago) some slight change may have taken place and the brain was reorganized to incorporate a language faculty. Maybe. That would be like angelfish stripes, polyhedral shells, etc. The understanding of the physical channel for natural selection is so limited that you really cannot have an opinion on this. You can make fun of it, if you like, or you can wave a banner about it. But that doesn't make much sense. It is simply not understood how the physical channel constrains and controls the process of selection, beyond simple cases. Lewontin is one who thinks that we'll never know the answer for human higher mental processes -- that by any method we can imagine now, there is no way to find the answer, not just for language but for cognition altogether. Others feel that they can do something. But telling stories is not very instructive. You can tell stories about insect wings, but it remains to discover how they evolved -- perhaps from protuberances that functioned as thermoregulators, according to one account. A famous case is giraffes' necks, that was the one case that was always referred to as the obvious example of natural selection with a clear function; giraffes get a little bit longer neck to reach the higher fruits, and they have offspring and so giraffes have long necks. It was recently discovered that this is apparently false. Giraffes don't use their necks for high feeding, end of that story. You have to figure out some other story: maybe sexual display like a peacock tail or some other story, but the point is that the story doesn't matter. You can tell very plausible stories in all sorts of cases but the truth is what it is. You can tell stories about the planets, as the Greeks did, in fact: nice stories, but things don't

work that way. In the case of language, we know that something emerged in an evolutionary process and there is no indication of any evolutionary change since it emerged. It emerged once, as far as we know, very recently. There is no real evidence for use of language prior to maybe 50,000 years ago or so. But the neuroanatomy seems to have been in place before that, so maybe 150,000 years ago. Anyway it's recent. The emergence seems to be fairly sudden, in evolutionary terms, in an organism with a very large brain, which was developed for whatever reason, and conceivably through some reconstruction of the brain that brought into play physical processes that led to something that works close to optimally, like a virus shell. If the minimalist thesis actually gains some significant credibility, that would be not an unreasonable conclusion; of course you have to establish the thesis.

**AB & LR:** So, language could have come to existence suddenly, through a single mutation, basically in its modern form, and natural selection wouldn't have had time to act on it. How can we substantiate this "evolutionary fable", as you call it in *Minimalist Inquiries*? What kind of evidence do we have of the recency of human language?

**NC:** Well, one thing is that there just weren't a lot of humans around, as far as anybody knows. Current estimates of the number of individuals... I can't reconstruct reliably from memory, but it may have been something like maybe 20,000 about a hundred thousand years ago, in fact a very small population, which then scattered widely. Unlike other large organisms, humans had escaped any limited ecological niche, so they were all over the place, presumably from a single origin. They were adapted to many environments. That means very small groups and not many of them. And then there was an increase...I mean, nothing like the explosion of the last couple of hundred years, but there was a substantial increase and that coincided roughly with the appearance of symbolic manifestations, various ceremonies and people buried with their tools, lots of things that indicate that there was complicated social organization. That's pretty hard to imagine without language. So that's the kind of evidence available. There is also some physiological evidence: Philip Lieberman has argued that their larynx sank [27]. Some scientists agree, some don't. Whatever it means, it is peripheral. On the perceptual side, there doesn't seem to be anything much detectable, and of course, as for thought systems, there are no records and not a great deal to learn from surviving nonhuman primates, so it appears.

### *V. Scope and perspectives*

**AB & LR:** In a recent lecture at the Scuola Normale of Pisa you quoted English 18<sup>th</sup> century chemist Joseph Black stressing the importance, for his discipline, of establishing a "body of doctrine" on the model of Newtonian physics. Generative grammar and, more specifically, the Principles and Parameters framework has certainly permitted many subtle and surprising discoveries over a broad domain, and one may argue that a significant "body of doctrine" on different aspects of human language has been established. Taking for granted the obvious fact that nothing is definitively acquired in empirical science, what are those aspects that you would consider "established results" in our field?

**NC:** My own view is that almost everything is subject to question, especially if you look at it from a minimalist perspective; about everything you look at, the question is: "Why is it there?" So, if you had asked me ten years ago, I would have said government is a unifying concept, X-bar theory is a unifying concept, the head parameter is an obvious parameter, ECP, etc., but now none of these looks obvious. X-bar theory, I think, is probably wrong, government maybe does not exist. If Kayne is correct the right parameterization is not the head parameter, but some other kinds of

parameters about optional movements, certainly plausible, possible. We just have to see...But I don't think that is so unusual. If you look at the history of the sciences, this is just the usual situation. Even in the advanced sciences almost everything is questionable...What I learned in college, let's say, in science courses, a lot of it would not be taught today. In fact, what was taught twenty years ago would be taught differently today in physics or chemistry. Some things are sort of stable. The Periodic Table is still there, but elementary particles are nothing like what we were taught... In fact in any live discipline you really don't expect the body of doctrine to be terribly stable. You'll get new perspectives, things will be reinterpreted. The changes often may not look very great from the outside but in a sense you can say the same thing about generative grammar for fifty years. From the outside it looks more or less the same, but from the inside you can see that it is very different and I suspect that that will continue. What are island conditions, for example? This has been a core topic of research for forty years now; I still don't think we understand that. There's certainly plenty of data that aren't understood; Paul Postal [28] has a recent book about it and I am sure that it has tons of data that don't work in any imaginable way. Such problems abound. And also, to my knowledge at least, there is no really principled account of many island conditions. On the other hand, something will remain stable. The difference between weak and strong islands looks stable; maybe we don't understand it, but there's something there that is stable. Also conditions on locality and successive cyclic movement look stable to me, at some level of abstraction. I strongly suspect that the difference between interpretable and uninterpretable features will turn out to be stable, though it is a recent observation, five years ago there was no discussion about it. In some fashion metrical theory will remain. Argument structure will also remain, as will properties of scope and reconstruction and the recent discoveries about fine structure. The essence of binding theory will remain, but probably will be reinterpreted. It's not that anything ever gets thrown out; the results about, say, ECP will remain but they may be parcelled out in different domains, maybe with different ways of looking at them and so on. But I don't feel that one can really predict much. It's a young field, changes are taking place fast, there's lots of unexplained things. I am sure there are new perspectives that we haven't thought of yet. I wouldn't expect, or even hope for stability. If there's stability, it means we are not going to get very far because, in the stage where we are now, there are just too many mysteries. So if the field remains stable, that means there are going to remain mysteries. That was true for chemistry at the time that Joseph Black wrote, the chemist you quoted, mid 18<sup>th</sup> century. Let's just consider what chemistry was like in the mid 18<sup>th</sup> century and what's like today. Black wouldn't be able to recognize the current discipline. In Black's days, it was still commonly assumed that the basic components of matter are earth, air, fire and water, that water can be transmuted into earth, and so on. Chemists had a substantial "body of doctrine" at that time, they knew a lot about chemical reactions, when they took place and how they took place, but the way of looking at them has totally changed. Take a look at Lavoisier for example, who sort of founded modern chemistry and created the nomenclature that everybody still uses -- and the nomenclature wasn't just terminology; it was supposed to be truth, it was designed to tell you the truth: so oxygen is the acid generator because that's its nature (which turns out to be false). In one of his classifications, alongside of Hydrogen and Oxygen we find "caloric", what we call "heat". So, everything has changed. And he kind of anticipated it; he said at that time that probably the nature of the elements is unknowable by humans, so we can just make some speculations. And chemistry was a pretty advanced science by that time.

**AB & LR:** Sometimes speaking with specialists of other disciplines, people ask "what are the results of modern linguistics?" Is there a way of phrasing some of the results, sort of independently from the technical language that makes them opaque for the public at large?

**NC:** There are things understood which you can illustrate easily, like, say, properties of *wh*-movement, which are very dramatic and a lot of them we understand at some level, for example Huang's distinctions and island effects [29], or even more complicated things like parasitic gaps

and so on. Even very simple examples can illustrate quite complex points. Sometimes I use examples like complex adjectival constructions (English is good for this, better than other languages with complex adjectivals), which illustrate successive cyclic movement in the predicate phrase, even though there's nothing visible, there is an empty operator. But the facts are clear and you can see the same facts that you see in *wh*- questions; you can state the principles that yield the interpretive facts in "John is too stubborn to talk to", that sort of thing. There's plenty of material like that, which is stable, easy to illustrate; you can state the principles, something that is known about the general principles. The fact that there is a component that deals with phrase structure in some fashion and a component that deals with dislocation in some fashion, that's, I think, pretty clear, and also that they have different properties, different semantic properties, different formal properties. The same if you move to phonology. So, sure there is a substantial body of things that can be presented in public talks; say, anything from middle school students to college and general public audiences. It is pretty easy to bring this kind of material to them -- and I am sure you do the same -- to get them to understand and even see the underlying principles. So, there are many nontrivial answers. On the other hand, if you ask for an axiomatic system, there is no such thing, but then you can't do it for any other science either. I mean, if somebody asks you what are the results of biology, all you can do is give an organized system involving natural selection, genes, Mendel's results and modern genetics, and so on, and then you can illustrate things...

**AB & LR:** The Minimalist Program has led researchers to rethink the foundations of their work, thus offering fresh perspectives on old problems, opening new questions, etc. On the other hand, the Program selects its own empirical domain on the basis of its stringent criteria, thus leaving out of its scope a significant part of the previously constituted "body of doctrine". Is this inevitable? Do you think it is desirable?

**NC:** It would be nice to subject everything to a minimalist critique, but it is quite hard; because nothing resists that critique, in any domain. So as soon as you look at anything, the best established work, and you ask, "Can I explain this just on the basis of legibility at the interface?" Well, the answer is no. That is true for the most elementary things, like sound-meaning correspondence: that's the basic data that people use, this sound corresponds to this meaning, that is everybody's basic descriptive data. But that doesn't satisfy minimalist criteria, the stringent ones at least. A stringent minimalist criterion would say: "The expression has to be legible on the sound side and has to be legible on the meaning side; but if it pairs up properly, that's something to be explained." You are not given that datum and that would require a much richer set of conditions imposed from the outside; in fact I don't think it would even be stateable as a set of conditions from the outside because in order to know that the pairing is correct, you have to know pretty much everything. So, somehow, even that simple datum, which every linguist for thousands of years has taken as the basic datum of the field, isn't available on a minimalist account. You have to try to explain it, you have to show that the optimal solution to legibility on the sound side and on the meaning side independently is going to give you the right interpretation for "John is easy to please", not some simpler interpretation. So the thing to do, at least it seems to me, is to pick the core pieces, like, say, phrase structure and dislocation, and ask what components of these systems look as if they are problematic. For instance, using the criterion that I think you had suggested earlier: would it be in an invented symbolic system? That's a good starting point. If you find something that wouldn't be in an invented symbolic system, you have to ask why it is in language: morphology, for example, what is it doing? And as soon as you ask, that drives you to new things, like the difference, for example, between interpretable and uninterpretable features, which is sort of transparent but I had never thought about it before, at least. It never occurred to me that there is a reason for the traditional asymmetry of agreement that we all learned in school. If you look at it from the point of view of ten years ago, I would have said the relation is symmetric and the traditional asymmetry is just arbitrary convention. But it is clearly not irrational, it is an intuitive perception of something

that appears to be quite deep; the distinction between interpretability in one position and not in another position. So, it's not trivial, but these things don't occur to you until you start asking: "Why is it there?" But then that proceeds for everything; everything that fell under ECP, under binding, under government, under proliferation of inflectional categories, almost everything. As soon as you begin to ask the simplest question, I think that the descriptions that looked obvious appear quite problematic, and the questions begin to proliferate as soon as you investigate. That's true of just about any point that you look at. In anything that you look at, you see that the assumptions are OK at some level and in fact revealing, some are very revealing, but then you start looking at the assumptions on which they are based, and find that the assumptions are dubious, they are not self-evident and sometimes not even natural. In particular, they surely don't follow from just the fact that the language has to be legible at the interface. Therefore you have to seek some other explanation for them and either you say: "Well, I give up, explanations have to end somewhere, it's a mystery," or else you look for an explanation, and the assumptions often dissolve. Anyway, we shouldn't accept the idea that it's a mystery. Maybe it is, but it's way too early to assume that. That is an admission of defeat that is surely premature. It could turn out to be right, maybe it's a mystery. We have been all along (and justly, I don't criticize this) willing to accept principles because they yield results. That's the right way to proceed, without asking why such principles exist. However at some stage, maybe it is too early, but at some stage it is going to be necessary to ask why the principles exist and a minimalist approach gives one way of looking at this. Maybe there's some other way...but I can't think of any other way at the moment.

**AB & LR:** One can address the same problem of empirical coverage from a slightly different perspective. On the one hand, the MP relies heavily on a theory of the interfaces, which should provide the external constraints to be met by the language faculty. As such, MP should promote research on the neighboring systems and the interfaces even more than previous models. On the other hand, the program so far does not offer much guidance for the study of systems that are assumed to be language-related, but differently constituted from "narrow syntax", in your sense. Do you think this is a contingency of the current state of research, and things could, or should change in the future?

**NC:** First of all, the focus on the interfaces is extremely recent; until now, it has always been assumed, as far as I know, without any question, that there are two interfaces. This goes back to Aristotle: there's a sound and a meaning, and that's it. You look at sound-meaning correspondences, phonetics tells you the sound, nobody knows what tells you the meaning. That has been the general assumption; and it didn't matter much. Whether the assumption was right or wrong it had no effect on the theories, because they were not designed to satisfy the interface conditions. As soon as you think about that, about the fact that the essential property of language must be that it satisfies the interface conditions -- and that much, everybody has to accept --, then the question arises: what are the interfaces? It didn't really arise before but now it's going to matter. As soon as you look at it, you see that we really don't know.

So, let's take the easy case: the sensorimotor interface. It has always been assumed that there is one, but that is not in the least obvious. There might be different ones for articulation and perception, and furthermore it is not obvious that there is one interface for either articulation and perception. Suppose that something like Morris Halle's picture is correct [30]: the features at some level are giving instructions to the articulators. Well they don't have to all do it at the same point in the derivation. Perhaps some give instructions at one point and then there could be more phonological computation, then another instruction is given, and so on. It could be a distributed system in this sense. That is possible. I mean, why should biology be set up so that there is one fixed point in the computation at which you have an interface? Interpretation could be "on line" and cyclic, and even at each stage of the cycle, instruction to articulators and the perceptual apparatus might be distinct in character (rather than a single phonetic representation) and distributed within

the computation. There might also be the kinds of interaction proposed in the motor theory of perception. These may involve interactions between two aspects of the phonetic interface. So, I suspect that there may well be all sorts of surprises.

On the other side, the meaning side, it seems to me that there may be some suggestive results. A lot of the most interesting syntactic work that is now being done (usually called “semantics” though it should be considered the edge of the syntax, I think) doesn’t satisfy natural minimalist conditions on the language faculty: binding theory, quantifier scope or even operations that appear to involve movement, like Antecedent Contained Deletion. These do not easily fit in the whole picture. For one thing the operations are countercyclic, or, if cyclic, involve much more complex rules transferring structures to the phonological component, and other complications to account for lack of interaction with core syntactic rules. It is conceivable that these are just the interpretive systems on the meaning side, the analogue to articulatory and acoustic phonetics, what is going on right outside the language faculty. Nobody really has much of an idea about the computational processes right outside the language faculty. One could say there is a language of thought or something like that, there are concepts, etc., but there has never been any structure to the system outside the language faculty. Well, maybe this is the beginning of discovery of some structure right at the edge, using operations similar to internal operations but probably not the same. They have different properties.

There are some interesting possibilities; for instance these operations on the outside don’t iterate. So, it seems you don’t have successive cyclic QR, successive cyclic Antecedent Contained Deletion. That is also true apparently of the operations that probably are on the sound side, between the internal syntax-phonology interface and the external interface between the language faculty and the sensorimotor system. Things that involve heaviness, let’s say, Heavy NP Shift, all the operations that fall under Ross’ Right Roof Constraint. These also don’t iterate. That part of the internal syntax is, in a way, peripheral. It is not part of what one would imagine to be the essential core of language: the mechanisms for formulating thought in internal linguistic expressions. The operations of the phonological component, broadly construed, are forced by the needs of the sensorimotor system. And if these operations have properties similar to those external to the other interface, that’s suggestive. So, maybe that is the beginning of some kind of nontrivial study of thought systems, how they are working at the point near the language faculty where you can gain at least some access to them... Those are new questions, questions that immediately flow from the insistence -- right or wrong -- that the internal operations have highly systematic minimalist properties.

The general point is that -- as is normal in the sciences -- you are trying to show how the language faculty meets certain conditions, but you have to discover those conditions, and you expect to discover what the conditions are in the course of the process of asking how the language faculty satisfies them. It’s not like the case of an engineer who is given the conditions and is told: “OK, satisfy them”. Here we are in a process of discovery, we have to find out what the conditions are and finding out what the conditions are is part of the process of finding out how to satisfy them, so the two processes are going to go hand in hand. If this whole approach turns out to make any sense as a research topic, it should lead to a much more careful exploration of the interfaces themselves, what’s on the other side of them. That should be a major research endeavor, which really hasn’t had much of a place in the subject until now.

Actually, the imaging work may be of particular interest here. Imaging studies should be particularly valuable in sketching out the general architecture of systems and how they interact, hence in exploring the ways in which the language faculty (or the several language faculties, if that is how the picture develops) interact with other systems of the mind/brain. Some light is shed on these questions by “nature’s” experiments” (brain damage, etc.), but direct invasive experimentation is of course excluded. The newly emerging technologies should provide a way to overcome some of the barriers imposed by ethical considerations and the diffuse effects of natural events. Even in early exploratory stages, there are results that are quite suggestive, and it may be

possible to design experimental programs that would yield important new kinds of information about the nature of the language faculty and the way it is accessed and used.

#### ACKNOWLEDGMENT

We would like to thank Marco Nicolis and Manola Salustri for editorial assistance.

#### NOTES

[1] Noam Chomsky (1981) *Lectures on Government and Binding*, Foris Publications, Dordrecht.

[2] See, for instance, the articles collected in Noam Chomsky (1977) *Essays on Form and Interpretation*, North Holland, New York.

[3] Noam Chomsky (1995) *The Minimalist Program*, MIT Press, Cambridge, Mass. See also Juan Uriagereka (1998) *Rhyme and Reason*, MIT Press, Cambridge, Mass., for an introduction to the basic concepts and empirical results of minimalism.

[4] Steven Weinberg, "The Forces of Nature", *Bulletin of the American Society of Arts and Sciences*, 1976, 29, 4: 28-29.

[5] Roman Jakobson (1936) "Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutung der russischen Kasus, TCLP, VI, English translation in Roman Jakobson, *Russian and Slavic Grammar*, Mouton, Berlin, 1984.

[6] Dianne Jonas (1996), *Clause Structure and Verb Syntax in Scandinavian and English*, PhD dissertation, Harvard University; Adriana Belletti & Luigi Rizzi (1988) "Psych-verbs and Theta Theory", *Natural Language and Linguistic Theory*, 6, 291-352.

[7] Ken Hale & Samuel Jay Keyser (1993) "On Argument Structure and the Lexical Expression of Syntactic Relations", in Ken Hale and Samuel Jay Keyser, eds., *The View from Building 20*, MIT Press, Cambridge, Mass.; Luigi Rizzi (1997) "The Fine Structure of the Left Periphery", in Liliane Haegeman, ed. *Elements of Grammar*, Kluwer, Dordrecht, 281-337.

[8] Guglielmo Cinque (1999) *Adverbs and Functional Heads – A Cross-linguistic Perspective*, Oxford University Press, New York, Oxford.

[9] See the discussion in Jerry Fodor, Thomas Bever, Merrill Garrett (1974) *The Psychology of Language*, McGraw-Hill, New York.

[10] W. V. O. Quine (1972) "Methodological Reflections on Current Linguistic Theory", in Donald Davidson & Gilbert Harman, eds., *Semantics of Natural Language*, Humanities Press, New York, 1972; W.V.O Quine (1986) "Reply to Gilbert H. Harman", in Edward Hahn and Paul Arthur Schilpp, eds., *The Philosophy of W. V. Quine*, Open Court, La Salle 1986.

[11] Dianne Jonas (op. cit.).

[12] Noam Chomsky (1998) "Minimalist Inquiries: The Framework", ms., MIT.

[13] Noam Chomsky (1964) "Current Issues in Linguistic Theory", in Jerry A. Fodor & Jerrold J. Katz, eds., *The Structure of Language*, Prentice-Hall, Englewood Cliffs, 50-118.

- [14] John Robert Ross (1967) *Constraints on Variables in Syntax*, PhD dissertation, MIT; Emmon Bach (1971) "Questions", *Linguistic Inquiry*, 2, 153-167.
- [15] Richard Kayne (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, Mass. See also the discussion in Chomsky (1995) op. cit.
- [16] On government and ECP see Chomsky (1981) op. cit. and much subsequent work.
- [17] Marc D. Hauser (1996) *The Evolution of Communication*, MIT Press, Cambridge, Mass.
- [18] François Jacob (1981) *Le jeu des possibles*, Fayard, Paris.
- [19] Editorial comments in Martin Joos, ed. (1957) *Readings in Linguistics*, American Council of Learned Societies, Washington.
- [20] Edward Sapir (1921) *Language*, Harcourt Brace, New York.
- [21] W. D'Arcy Thompson (1917) *On Growth and Form*, Cambridge University Press, Cambridge.
- [22] Johann Wolfgang Goethe, *Bildung und Umbildung organischer Naturen*, 1807.
- [23] See Alan Turing's classical paper "The chemical basis of morphogenesis", *Philosophical Transactions of the Royal Society of London*, 1952, 37-72, and the review of the issue in Ian Stewart (1998) *Life's Other Secret*, John Wiley, New York.
- [24] Walter J. Gehring & Kazuko Ikeo, *Trends in Genetics*, Sept. 1999.
- [25] Jerry Fodor (2000) *The Mind doesn't Work that way: Scope and Limits of Computational Psychology*, MIT Press, Cambridge, Mass.
- [26] Richard C. Lewontin (1990) "The evolution of cognition", in D. Osherson, E. Smith, eds., (1990), *An Invitation to Cognitive Science: Thinking*, vol 3, MIT Press, Cambridge, Mass., 229-246.
- [27] Philip Lieberman (1984) *The Biology and Evolution of Language*, Harvard University Press, Cambridge, Mass.
- [28] Paul Postal (1999) *Three Investigations of Extraction*, MIT Press, Cambridge, Mass.
- [29] James Huang (1982) *Logical Relations in Chinese and the Theory of Grammar*, PhD dissertation, MIT. On parasitic gaps see Noam Chomsky (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge, Mass. and references cited there.
- [30] Morris Halle & Kenneth N Stevens (1991) "Knowledge of Language and the Sounds of Speech", in Johan Sundberg, Lennart Nord, Rolf Carlson, eds., *Music, Language, Speech and Brain*, Macmillan, London, 1-19; Morris Halle (1995) "Feature Geometry and Feature Spreading", *Linguistic Inquiry*, 26, 1-46.