Studies in Chinese Linguistics
中國語文研究

Poverty of Stimulus: Unfinished Business*

Noam Chomsky
Massachusetts Institute of Technology

What I would like to do is to make a few remarks about the conceptual foundations of much linguistic work, and much contemporary work, and also on some interesting, and sometimes far-reaching consequences, that I don't think receive sufficient attention. And as a kind of paradigm example for research, I will turn to one of the simplest of the enumerable poverty of stimulus problems that immediately came to light as this approach to language began to take shape in the early 1950s. A very simple solution was suggested at the time. I think it is basically correct, but some fundamental inadequacies have been overlooked right to the present, and investigation of these opens up some rather difficult problems that are right at the borders of inquiry today. That's the specific example of unfinished business that I intended in the title. But it is important to emphasize that what is unfinished or barely even contemplated goes vastly beyond.

There are many ways to study language. There's one way that has a certain logical priority in the sense that every approach to language presupposes its conclusions, tacitly at least. That includes approaches that vigorously deny doing so. That's the study of the linguistic capacity of the user of language. Every other approach must be based on assumptions about these capacities, and they're usually tacit. That's true whether it's the study of the use of language in social interchange, or historical and comparative studies, or for that matter even something as simple as transcribing a corpus of material. Such assumptions are buried in somewhere. Thinking of language this way, it is an internal property of an individual, so sometimes these days called "I-language".

Uncontroversially, I-language at its core is a system of discrete infinity like the natural numbers, and any approach to language that doesn't deal with that fact is missing a core feature of language. The study of I-language is based on some method for characterizing an infinite discrete set of structured expressions. The

---

* Editor's note: The original version of this paper was transcribed by Professor Fuzhen Si, with light editing by Professor Riny Huijbregts, Professor C.-T. James Huang, and Professor Noam Chomsky, and published in *Universals and Variation: Proceedings of GLOW in Asia VIII 2010* (Beijing: Beijing Language and Culture University Press, 1-17) edited by Mingle Gao. This printing is by permission of both Professor Noam Chomsky and courtesy of the Beijing Language and Culture University Press and the volume's editor.

Studies in Chinese Linguistics, Volume 33, Number 1, 2012, 3-16

standard way to do that is by a generative procedure. There are alternatives, and maybe I'll say a couple of words about them later. I assume that the proper way is a generative procedure, and I think there is evidence for that. Well, we can then understand I-language to be nothing but the generative procedure that determines an infinite array of structured expressions, hierarchically structured expressions, each of which you can think of as a set of instructions to some of the other internal systems of the mind and body – instructions to the thought systems, sometimes called the "conceptual-intentional interface", and to the sensory motor system, the second interface. That's basically what an I-language is. Looked at this way, I-language accommodates a traditional characterization of language that goes back to Aristotle who defined language as sound with meaning. I will come back to that. It is basically correct, but I think it needs a fundamental revision, to which I'll return.

We're interested in I-language not just in what's technically called "extension" – that is, the class of generated structures – but rather in what's called "intension", that is the actual formulation of the "algorithm", not some other formulation that has the same extension. So you want to discover the intensional characterization of I-language, and this should be a central goal of the study of language. That's a fact that's been pretty seriously misunderstood, particularly in computational cognitive science and in philosophical literature, but is important to bear in mind. Well, at least that is our interest if we think of I-language as a real entity in the biological world, in essence, what's informally called an "organ of the body", rather like the visual system, the immune system or the digestive system and others. The "organ of the body" is just some sub-component of the complex organism which has sufficient internal integrity so that it makes sense to study it in abstraction from the rest of the system in which it participates, and then also to look into how it integrates with other similar sub-systems of the body, often called "modules" in the biological literature. In this case, I-language is a cognitive organ, rather like systems of planning and interpretation, reflection, and whatever else falls among those aspects of the world that are termed "mental," which reduce somehow to the organical structure of the brain. I am now quoting the very important philosopher and scientist Joseph Priestley in the late 18th century, who was summarizing the common understanding of post-Newtonian Europe when it was recognized (but has since often been forgotten) that you can't make a distinction between body and mind because Newton had eliminated the concept of body. So these are just different aspects of the world, some of them are termed "mental".

This approach to language began to take shape in the early 1950s. It was part of what some called the Cognitive Revolution, which departed quite sharply from prevailing conceptions of structuralist and behavioral science approaches to human mental activities and social organization, and adopted the approach of normal biology. That was a sharp break at the time, later coming to be called the "biolinguistic program". From this point of view, several questions immediately arise. Loosely, we could call them "what" questions, "how" questions and "why" questions. So "what" is the nature of the system of the I-language? "How" is it

acquired through its development by a child? And "why" does it have these properties, not other properties? Those are the kinds of questions asked about any such system and, in particular, about I-language. In technical linguistics they've got names: "what" questions are called questions of "descriptive adequacy", the "how" questions are called questions of "explanatory adequacy". The "why" questions have no standard name: just "beyond explanatory adequacy."

Particularly in an emerging scientific endeavor like contemporary linguistics in the modern sense, it's a useful idea to reflect some on the origins of the modern sciences in general. The modern scientific revolution that started in the 17th century provides us with some useful guidelines. For example, it's useful to remember that the modern scientific revolution began with a willingness to be puzzled about things that seemed entirely simple and obvious; let's say, with the fact that certain objects like apples fall to the ground if they're detached, while others like steam rise to the sky. For two thousand years, scientists had been satisfied with an answer to that: the Aristotelian doctrine that things move to their natural place. When 17th century scientists allowed themselves to be puzzled by such obvious facts, and asked why the world works this way, they very quickly found that there are quite serious problems. In fact, Galileo was able to refute common sense beliefs, which are still common sense beliefs, for example that the rate of fall is proportional to mass. When these questions were answered, partially at least, others quickly came into view, including many questions which are still at the borders of inquiry today. So for example, why gravity is a weak force – luckily for us, or we would be compressed to a point. Or for example such questions as "Where is 90% of the matter in the universe?" which no one can find, and is still a puzzle. It's just postulated, called "dark matter". It is postulated because otherwise theories don't work. A lot of work right now in Switzerland particularly is trying to see if you can maybe find an answer to this rather small difficulty. It's also worth recalling another leading theme of the modern scientific revolution, that's Galileo's thesis, which he was perhaps the first to enunciate very clearly, the thesis that nature is basically simple, and the task of the sciences, what defines them, is the effort to demonstrate that fact, which is pretty hard. That typically requires putting recalcitrant data to the side, like "where is 90% of the universe", putting it to the side and continuing to hope that with better understanding the phenomena will fall into place from some different point of view that we don't know how to contemplate now. We know how rich and profound is the legacy that grew from such small beginnings, and it's good to keep that in mind, as a lesson for today.

Turning to our topic, language, let's begin with the "How questions", questions of language acquisition. Every aspect of growth and development (language included), growth of arms and wings, anything you look at, has to involve at least three factors. I stress "at least" because there are others. It has to involve at least the data that's available which affect the nature of the growth and development. Secondly, we now know, the genetic endowment. And thirdly, laws of biology, or perhaps more generally laws of nature that are involved somehow

in this process. In principle, these can interact in quite complex ways. So, for example, the course of growth is known to be able to modify gene expression. And in turn, that could mean that universal grammar – the genetic endowment – could be different for an adult and a child. Unless there's some evidence to show that this happens, it doesn't make much sense to contemplate these possibilities, although we should keep them in the back of our minds. There's another factor that enters into growth and development that's truly significant: that is the structure of the brain. The structure of the brain only allows certain courses of growth and development of cognitive organs and not others. Here again, essentially nothing is known, too little to be able to bring this consideration into play when we study the "how questions". And that's understood. So for example, the distinguished neuroscientist, Eric Kandel wrote a recent review of the field. He writes that the neuroscience of higher cognitive processes is only beginning. And in fact even that might be too optimistic. There's a very important recent (co-authored) book by Randy Gallistel, a distinguished cognitive neuroscientist. He makes quite an extreme, but I think pretty plausible argument. He argues that throughout their entire history, that's hundreds of years, the neurosciences of higher cognitive processes have been looking in the wrong place. The neurosciences are seeking answers in a place where they're never going to find them and therefore the field has to just be reconsidered from its roots. What he argues basically is that the neurosciences have been misled by what amount to associationist dogmas, and are therefore looking at such elements as strength of synaptic connections. His argument is that these just can't, in principle, yield the properties of memory and other cognitive functions even for insects, let alone human language. And that in fact what they should be doing is starting by asking "what are the cognitive systems that we are trying to discover the basis for". So let's look at the properties of, say, bee navigation and human language, and then try to see if you can find the elements that are involved in those. These are some kind of computational system. Therefore, the goal of the neurosciences should be to find the neural basis for the elementary units that enter into any computational system, units that have properties like *read*, *write* and *address*, familiar from Turing Machine Theory, any form of computation. Well, I think he makes a plausible argument, though it has been kind of disregarded. It is a pretty radical conclusion, but I suspect it won't be disregarded for long. In any event, right now at least, the brain sciences just don't provide much relevant information for the "how questions", so we have to keep to the factors that do provide some information, keeping in mind however, that we might be missing something important that has yet to be discovered and that might be related to the inadequacies of attempts of explanation.

Among the genetic factors, we can make a further distinction between those that have to do with cognitive capacities outside of language, maybe they're general, maybe they're specific, but anyhow outside of language; and others that are specific to language. In current terminology, the genetic factors specific to

language are called Universal Grammar (UG). It's an old term adapted to a new usage. It probably shouldn't have been used because it's misled all sorts of people, but anyhow that's the standard term. UG is just the theory of the genetic component for language, whatever it is.

Just what these factors are and how they interact in language acquisition is a very live research topic. The most promising approach I know of is by Charles Yang, who should be here I think, in his book, *Knowledge and Learning in Natural Language*. It can hardly be doubted that both components exist and that they participate in answering the "how" questions. So for example, if UG did not exist, it would simply be a miracle that every infant somehow selects out of the environment, just a confused environment, data that are relevant to language, reflexively. And then the infant also more or less reflexively attains the language mastery that all of us have. Meanwhile, other organisms faced with exactly the same data can't even take the first step, and obviously none of the later ones. This is a rather simple and obvious fact that's consistently denied in related fields, which is an indication, I think, of their kind of pre-Galilean status on willingness to face elementary facts of the world. There are other kinds of evidence that point in the same direction, like dissociations between language and cognitive capacities, but the basic conclusions seem clear. More generally, as Randy Gallistel points out from a computational point of view, the notion of a general purpose learning system makes no more sense than the notion of a general-purpose sensing organ. Like a bump in the middle of the forehead whose function is to sense things. What has regularly been discovered, he points out, from insects to humans, is modular systems, with very special growth learning mechanisms in different domains among different species. Now there is every reason to expect that language keeps to the biological norm in this respect. There are crucial features of human language that appear to be isolated in the biological world. They seem to have developed very recently in evolutionary time, many millions of years after the separation of humans from any other surviving species.

Let's turn to the "Why questions". These lead to the search for general principles that enter into answering the "what" and the "how" questions. Well, since language is clearly a computational system, one factor that you want to look at is computational efficiency, probably a law of nature. We don't have a comprehensive theory of computational efficiency, but quite a few things are pretty obvious that are going to enter into such a theory. One, for example, that less is better than more. So a rule system, X, let's say, is preferable to X+Y (or Ys, some additional rules); or for example, that a minimal search is more efficient than deeper search. These principles of minimal computation will surely enter into any ultimate biological law or maybe a more general theory of computational efficiency. That assumption takes on special salience in a case like ours, the case of language, because of some obvious observations. There's very close uniformity in outcomes with varying input data in language acquisition. That indicates there must be very strong uniform constraints, either UG constraints or such third factor properties as minimal computation.

Well, in every domain of growth, including acquisition of a particular language, there's an enormous gap between the data that are available and the state that's attained. Another truism that's consistently denied in the case of language, though it is understood everywhere else. This universal property of growth and acquisition has been given a name, but just in the case of language. In other domains, it's considered so trivial, that nobody even bothers giving the name. In language it's called the problem of "poverty of stimulus." That's considered highly debatable or radical, but is close to truism. It's interesting to discover exactly what it is. It's always going to be there and huge. This problem has to be overcome by the interaction of other factors, including genetic factors and the third factor, biological and more general natural law.

In the case of language, we can also formulate a narrower "poverty of stimulus" problem. Namely, we can ask, "what is the contribution of one of the parts of the genetic component", namely, universal grammar. Well, investigation of these topics has generally kept to syntax and phonology, with kind of barely a nod to semantics. That's arguably practical necessity today, given inadequacy of understanding. But it should be realized that in the domain of semantics, the problems of poverty of stimulus, and also evolutionary origin (they are related), those problems are utterly overwhelming. So if we are willing to be puzzled by puzzling facts, then we find very quickly that the evidence for the semantic interpretation of expressions is very slim. That's true even for the most elementary meaning-bearing elements, let's say, words, for simplicity. Those problems are usually evaded by the referentialist dogma, which holds that there is direct relation, maybe even a causal relation, between words and extra-mental entities. That's why standard texts have titles like "words and things" (Roger Brown) or "word and object" (W. V. Quine). The assumption actually appears to be correct for animal communication systems, but it's radically false for human language – an important fact that's often overlooked. The problems of answering such simple questions as this, "what's the relation between words and extra mental entities", are awesome. And there's a comparable problem for the evolution of human language or perhaps evolution of human thought. The topics have barely been addressed mainly because they haven't been recognized. And in this domain I think it's fair to say that the fields are in the state of neo-scholastic physics pre-Galileo, satisfied that "things move to their natural place".

Keeping to syntax and its externalization to the sensory motor interface, it's very natural to attribute a rich and complex structure to these processes. The reasons are just practical necessity, but it should be regarded again as a temporary expedient. In fact, for theoretical linguistics, the central task from the beginning of the 1950s has been to try to show how to reduce complexities of universal grammar. In recent years, this effort has been given a name. It's called the Minimalist Program, but that's misleading, too, because everything that's done is the minimalist program. In fact that's just a name for science. The reasons for trying to simplify the UG are several. One is just standard science: we hope to find deeper explanations, and that

means to eliminate stipulation. Another reason is the hope that eventually there may be something to say about evolution of the language capacity. Evidently, the richer the assumptions about universal grammar, the harder that extremely difficult task is going to be.

These efforts have proceeded quite far. It's perhaps worth noting that extensive efforts to reduce the complexity of UG by statistical analysis of corpora, or similar methods of corpus analysis, efforts that dominate almost all of computational cognitive science, have been quite barren by comparison. It's a topic that merits some serious thought, I think.

For syntax, the earliest proposals about UG in the 1950s had a kind of traditional flavor. They that assumed phrase structure rules that yield hierarchically structured expressions and transformational rules that convert these to the surface form, and also that deal with the ubiquitous properties of displacement in natural language, the fact that expressions are commonly pronounced in one position but interpreted not only there but also in other positions. For example, the English sentence "which book did you read?" is interpreted more or less as "for which book x, you read book x". The phrase "which book" is interpreted in two positions: in one position to receive its semantic role as the object of "read", in the other position to be interpreted as a part of an operator-variable structure.

Descriptive evidence from a wide range of languages very quickly seemed to show that each of these rule systems was very complex. And that the manner of their interaction was intricate and varied widely from language to language. That's a paradoxical, in fact impossible conclusion because of the simple fact that languages are acquired very rapidly and in uniform ways, and furthermore, that uncontroversially humans are not genetically predisposed to learn one rather than another language. So from those observations it follows that these conclusions have to be wrong, even though there didn't at first seem to be any alternatives.

In the 1960s, it was shown that the very complex stipulations of phrase structure grammar could be substantially eliminated in favor of X-bar theory. Then and in later years, the study of conditions on transformations allowed quite sharp reduction of the postulated complexity of the transformational component. What remained as a primary conceptual problem was to try to reduce these two systems to the same operation; if possible, the most elementary operation. By about ten years ago, I think, it was pretty clear that that looks possible, along lines to which I'll turn in a moment. But first I would like to mention that the description is misleading. So just take the phrase structure component. The reduction to X-bar theory is substantial, but the further reduction of X-bar theory to simpler terms that has taken place isn't actually a reduction to further simplicity, It's a matter of teasing out various things that were put together in X-bar theory but that should be separated. One is the property of forming expressions by combining parts, another is linear order, and a third is the matter of identifying what kind of expression they are. X-bar theory mixes these together. But they're really separate, and each of

them, which when looked at, raises quite significant problems which are again at the border of research. Well, I put that aside, though it is interesting.

When the biolinguistic program began to take shape in the 1950s, there was a general assumption about language, namely that the basic problems of language and of psychology generally had been solved, or at least were very close to solution in terms of structuralist procedures and the methods of the behavioral sciences, which were expected to become more effective because of new conceptual tools like the mathematical theory of communication, and also with the availability of computers, which were coming on line at the time. As soon as investigators allowed themselves to be puzzled by simple phenomena, it quickly became clear that the assumption was completely wrong, that almost nothing was understood, and that everything when looked at posed very serious problems. In that respect the situation resembled pre-Galilean physics. Many examples were quickly proposed as illustrations. One particularly elementary one has taken on a kind of classic role, though it's only one of many. It's a problem of the formation of simple "yes-no" questions in English or similar issues in other languages. Failure to be puzzled about that was a big mistake, which was quickly found, and failure to be sufficiently puzzled about it raises questions that are still at the borders of research, and that leads in very strange directions, if you look at them carefully.

So just taking this simple example, I think, sheds a good bit of light on the general enterprise of biolinguistics. Take, say, the sentence "Young children can write stories". There's an interrogative counterpart "Can young children write stories?". What's happening is that the preposed element "can", verbal-like element, is clearly related to the word "write". That's obvious from the interpretation and even obvious from morphology when there's visible reflection like "are young children writing stories" and so on. That's of course only one example of a very general phenomenon, in English and cross-linguistically.

What about the "how" and the "why" questions? They arose quite sharply when the problems were posed about fifty years ago. So suppose you expand the sentence to "Can young children who work hard write stories" with the two verbs "work" and "write". Both the interpretive and morphological evidence show that the preposed inflectional element is associated with "write", not with "work". It's not in question, but there is a question as to why it should be so. (Now we are back to "why objects fall to their natural place"). It's not so obvious. From a computational point of view, it would be easier to raise the element corresponding to the first verb, not to the second verb. It is much easier to find, and easier to parse, and it doesn't require a complex structure, just linear order. In terms of acquisition, why doesn't the child do the simplest thing in acquisition. It never happens in English or any other language that a child ever makes that kind of mistake. So something serious must be going on. And there was an answer proposed. The answer was that rules are what was called "structure-dependent", that is, the distance is measured

structurally, not linearly: presumably a UG property. There is also an implicit recourse to the third factor principle of minimal computation: you try to do things in the simplest way and you find the closest element permitted by UG.

It might be worth mentioning that this elementary illustration of the poverty of stimulus problem has taken on a life of its own. A major topic of computational cognitive science is to show that that's not right, and that you can somehow get the right result by a complex enough statistical analysis of corpora. There's a huge literature on this, which is rather strange for several reasons. One reason is that there is a very simple answer, always ignored. The second reason is that every effort fails, and fails dramatically, so much so that there is no way to fix it up. Or else, they just beg the question by restating the conclusion, one or another way. Another reason why this is strange is that wouldn't matter much if any of the approaches worked, because they're not dealing with the question. So suppose one of these statistical approaches worked, the question immediately arises "Why don't half the languages in the world use the most prominent element, and the other half use the closest element?" You can get that answer from a corpus in an invented language by essentially the same methods. That was the original problem. So even if they worked, which is far beyond imagination, it wouldn't matter. But that doesn't change anything. Journals like *Cognitive Science* and others regularly publish failed approaches to the problem, without recognition of the fact that there's a simple answer, and it wouldn't matter much if it worked. So it's strange. It's a question for the sociology of science, I think.

There still are "how" and "why" questions, even if we accept the "what" answer. Why does it work? Well, minimal computation is presumably a third factor property; it's just a natural law. The best explanation for the choice of structural rather than linear distance would be that linear order is simply not available to the operations of the I-language – that it is a secondary phenomenon imposed by the sensory motor system. Now that makes conceptual sense. Remember that the I-language has to provide instructions both to the thought systems and sensory motor system. As far as we know, core semantics doesn't pay any attention to order, it just pays attention to hierarchy, and therefore there is no reason why the I-language should impose order apart from the demands of sensory-motor system, which of course do require order. That would be the natural result. The natural conclusion would be that word order is just a property of externalization to the sensory motor system, a secondary phenomenon in language, not accessible to the raising operation. Well if that were correct, we'd have a convincing and principled answer to the "how" and "why" questions. But there are problems. The problems are both empirical and conceptual. The empirical problem is that there seems to be evidence that linear order is involved in the computations that yield semantic interpretation, even though it's not used for such interpretations. It is kind of paradoxical, particularly because it's clear why the sensory motor system imposes the ordering requirement. At this point, research intuitions diverge. We're back to a familiar question, "do we assume that the data just compel us to accept hopelessly

implausible conclusions?" or "do we try to show that the data have been somehow misinterpreted?" There's another way of looking at it, and like 90 percent of the matter of the universe, we'll find a better answer down the road. Well, it's pointless to argue research intuitions. From my own view, for what it's worth, the latter one is the sensible one.

There's also a fundamental conceptual problem that has passed completely unnoticed; that's another case of insufficient puzzlement. Time's brief so I won't go into any detail, but if you think about the basic structure of the subject and predicate, it's generally assumed that the prominent element of the predicate is what's raised, but why? I mean the two parts are just on par; there's no difference between them. Why don't we raise the most prominent element of the subject? That would mean that the question corresponding to "Young children write stories" is "Children young write stories". Yes or no? Well, that's obviously inconceivable. But it does raise a question. And when you look at that question, when you permit yourself to be puzzled about it, it turns out to be quite non-trivial and there are very rich consequences. There are indefinitely many problems like these.

Let's try to take a more systematic look at the whole array of problems, beginning with Aristotle's common sense dictum that language is sound with meaning. In millennia of study, almost all the attention has been devoted to sound, more generally externalization, so phonology, morphology, linear order and so on. The focus on externalization is understandable. That's what you see, it's the locus of complexity in language and of historical variation, historical change; it's what you have to learn when you learn a language or what you try to teach if you teach language, and so on. But it might turn out that all of this is secondary. It's not really a language. But it seems that it is really an ancillary component relating language to some other system, the sensory motor system. That would mean that the Aristotelian dictum ought to be reversed: language is meaning with a sound, a very different concept. And not necessarily a sound, as we know. It may be a sign, it may be nothing external at all as when you're talking to yourself. That is the most widely used aspect of language. I think that's probably correct and the consequences are quite far reaching.

There has, of course, been work on meaning in recent years, quite a lot, and highly illuminating. But I think we should understand it as syntax; that is, it has to do with symbol manipulation. It doesn't matter what the world is. In that respect it's kind of like phonology. Nothing is wrong with phonology, but it doesn't deal with the mind-external world. There's a study that does deal with this world: that's called acoustic-articulatory phonetics, which is not phonology. The comparable field relating internalist semantics to the world remains unanswered, in fact unasked, primarily because of the force of the referentialist dogma. If it's false, if nothing replaces it, then these are studies of syntax, symbol manipulation, which is not a criticism any more than it's a criticism of phonology to say that it doesn't deal with

acoustic phonetics. In this respect, language does seem radically different from any animal systems known, which do appear to satisfy the referentialist dogma.

There's a third notion in the Aristotelian dictum that language is sound with meaning – or, as I believe it should be rephrased, meaning with sound. That is the notion *with*: how meaning and externalization, maybe sound, are linked. That's barely been studied in any depth until recently because the answer was taken to be obvious, and still is often taken to be obvious.

The core question has to do with novel structures, which occur all the time. These central properties of language have been known for a long time. Galileo was struck by them, and they are at the core of Cartesian science. But they received very little attention in the thousands of years of language study – reminiscent of the doctrine that "things fall to their natural place". When anyone asked the question, which was rare, the answer is induction or analogy. It still often is. These concepts are left unanalyzed. So that's pre-Galilean. You have to ask that question if you want to enter into serious scientific discussion of language. Language is again a system of discrete infinity like the number system, so it has to include some kind of combinatorial procedure. A combinatorial procedure at its root forms new objects from objects already formed. Let's call that operation Merge. In the best case, the ideal case, that would be the only combinatorial operation in I-language, maybe even the only principle of UG. The other properties of language would consist of nothing more than applying third factor properties like minimizing computation. There's a thesis to that effect: it's called the Strong Minimalist Thesis. It is an ideal – nobody knows how closely one can approach it. For now at least it's very far from achievable. Whatever the answer is to this question, the study of evolution of language must at least account for the appearance of Merge at some point in human evolution; rather recently in evolutionary time, perhaps roughly seventy five thousand years ago, it seems. Any other properties of UG would provide further burdens for some eventual theory of human evolution. There are libraries full of speculation on this topic – though in fact the speculations are about evolution of communication, plainly a different topic. Conferences, books and so on will be of little value unless they address the elementary question.

Putting aside a great deal, we would take for granted unless proven otherwise that the operation Merge is as simple as possible – it involves minimal computation. That means when you form something new from X and Y, you don't change X, Y, and you don't impose an order on them, so the output of Merge of X, Y is set containing X and Y, {X, Y}. That's the best case, and it makes sense. It also makes sense from an evolutionary point of view because whenever Merge arose, let's say, seventy five thousand years ago, by some mutation, it was subject to no selectional pressures whatsoever. So therefore, it would develop just like a snowflake, just by natural law. So from an evolutionary point of view, that's what you expect. An ideal would be to see how close we can come to what you expect. Notice also that Merge in this respect, just by logic, has two types. If X and Y are merged, then either one

is part of the other or they are disjoint (intersection is impossible under the simplest assumptions). External Merge is the case when they are disjoint: let's say, forming "write stories", from "write" and "stories". If one of them is inside the other, say, "John read which book", and you pick "which book" and you merge it to the whole expression, you get "which book John read which book". That's the right form for semantic interpretation: the interpretation, read off directly, is something like "for which book x John read the book x." But it's clearly the wrong form for externalization. That's a simple example of displacement. To block one or the other of the two kind of Merge would require stipulation, because they both come free. And it happens that though what comes free gives the right form for semantic interpretation. It's totally wrong for phonetic interpretation: you don't pronounce the copies; only the most prominent one (with an interesting class of exceptions that in fact support the principle). That follows from another application of the principle of minimizing computation. You want to minimize the amount that's externalized; it saves an enormous amount of computational effort. It happens that by minimizing computation, the outcome is a form that is appropriate for semantic interpretation, but causes enormous difficulties for communication, perception, and for parsing. Anybody who has worked on a parsing program knows that the main problem is to figure out what appears in the gaps. Problems wouldn't rise if you'd just pronounce them. So what we have is conflict between communicative efficiency and computational efficiency. And computational efficiency wins hands down dramatically. If you look carefully, the number of copies that appear, yielding interpretive possibilities, is quite large, a phenomenon called Reconstruction. And the number of copies erased yields an enormous saving in externalization. There are other examples of conflict between computational and communicative efficiency. There's no time to go into them, but everyone I know of is resolved conclusively in favor of computational efficiency, which indicates again that the Aristotelian doctrine was backwards: language is meaning with sound (or some other form of externalization), and the externalization part is ancillary to language.

That poses a hard cognitive problem: how do we relate this approximation to a snowflake to a sensory motor system that had been around for hundreds of thousands of years. That is a hard problem and that's why you get the complexity of language. These properties are essentially universal.

I'll just make one comment on the puzzle I mentioned in the beginning, the "why" question concerning formation of simple interrogative questions: why do you pick the inflectional elements to raise, not the head of the subject, the nominal expression? There's only one simple answer to that that I can see: the nominal expression just isn't there at the point where the relationship is established. It's added later on after the relationship is established, so the question doesn't arise. That means that it's got to be merged from somewhere else. That's a familiar hypothesis. It is called the "predicate internal subject hypothesis." It's been pretty widely assumed but without very much evidence. However, this seems to give

substantial evidence for it. That means the subject is merged internal to predicate, and raised to the subject position – but after the relationship between the initial element of the sentence, the complementizer, and the inflectional element has already established. A lot of consequences follow from that. And it seems correct. Notice that it looks like a counter cyclic operation, which it is, superficially. But if you look at it from the point of view of what's called "phase theory," the conclusion falls into place fairly naturally. If that's what happens, it then raises another question: why does the subject have to raise? That's an old question called EPP. The "E" is there because nobody knows why it should apply: it's the Extended Projection Principle. Here is at least a partial answer. The partial answer is that unless you raise the subject, you have another XP-YP construction, and you can't tell what it is, because there're two parts which are on a par. If you have a structure with an external argument and a verb phrase, a transitive verb phrase, then either the external argument has to raise, in which case it's the head of the verb phrase, its complement, that you can label. Or else, the internal argument has to raise in which case the subject is technically the complement of the verbal head so it can be labeled again. And in fact that seems to be a fairly general property of language, that either the subject or the object has to raise. That raises new questions. What happens if there is an indirect object? Then you get the same problem again. That means that traditional grammar must have been correct calling it an indirect object "indirect". It means it's merged differently from the internal object. Probably that means technically it's pair-merged off somewhere on a different dimension, kind of like an adjunct. That raises further questions, like, why does it behave like an adjunct? Now we're running off into a huge mass of problems that arise. If you think about it, these problems will arise dramatically for every case of internal merge, movement, because it always forms XP, YP structures. How can they be identified? That raises new questions.

Well, I'll just leave the questions. But the point I'm trying to make is that if we allow ourselves to be puzzled about very simple phenomena, like apples falling to the ground, it turns out that a mass of questions arise, reaching quite far, and a lot of them are unsolved and at the borders of study. There are a couple of general conclusions and general lessons that are worth deriving from this little exercise, I think. One is that that it makes good sense to accept the Galilean principle that we should allow ourselves to be puzzled by things that look obvious. If we do, all kinds of intriguing doors open. And in fact, many of the questions that arise are right at the borders of inquiry just as in physics for that matter. Second is that if we do permit ourselves to be puzzled, we find that Aristotelian dictum, of the last twenty-five hundred years, is not accurate: language is actually meaning with an externalization, which could be sound, may be nothing, or may be something else. The study of the externalization, of course perfectly legitimate and significant, seems to be ancillary and more complex, because it involves cognitive processes other than those of language, and even physiological questions like the nature of the sensory motor system. The third

and last lesson is that I think it does make sense to take seriously the Galilean concept that nature is really simple. And if it doesn't look so, it's our fault. Our problem is to show that that's really true.

Mailing address: Department of Linguistics and Philosophy,
Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A
Email: chomsky@mit.edu
Received: March 23, 2012
Accepted: March 23, 2012