Poverty of Stimulus: Unfinished Business

Noam Chomsky

Lecture presented in the Lecture Series „Sprache und Gehirn – Zur Sprachfähigkeit des Menschen" organized by Angela D. Friederici in the context of the Johannes Gutenberg endowed professorship summer 2010[*]

[Chomsky:]

I am going to adopt a point of view, which is basically, that language is a subsystem of the organism, in this case mainly the brain, an "organ" in the sense in which the notion is informally used, rather like the visual or immune systems. It's a point of view that began to take shape in the early 1950s among a very small number of graduate students at Harvard - two or three - who were quite dissatisfied with the reigning doctrines of the day. In psychology, the doctrines were behaviorist and, in fact, radical behaviorist, and in the social sciences, they were what we call behavioral science, meaning you study behavior not what lies behind it. In linguistics, it was structuralism - either American or European - which had a kind of a similar cast. It was mostly analysis of data, what's called a corpus, with various techniques. It felt to us - we were a few mavericks there – that that's a very odd way of looking at human behavior. It's as if physics was called meter reading science. Something's wrong with that - it's true that the data may come from reading meters, but that's not what the science is about. It's trying to find out what's the reality of the world, what's the internal hidden structure of things, why do things happen the way they do, and physics itself, modern physics, really began with a willingness to be surprised at things that seem perfectly obvious. So if you go back to *Galileo* and his successors, the early days of modern science, there were a lot of things that were just taken for granted in scholastic neo-Aristotelian science. For example, if there's an apple on a tree and the branch breaks, the apple falls to the ground. Why? Well that's its natural place – that's its natural place, where it's trying to go. And that was considered - it's the common sense answer -- but *Galileo* and other scientists decided to be puzzled about that. So why does it fall to the ground instead of falling up? It was not easy in those days to convince the equivalent of the National Science Foundation, the aristocrats, that there was some point in studying these things. Why should they get support for studying, say, what

happens when a ball rolls down and a frictionless plane - which doesn't exist? What's the point of studying that, when you could be studying things are happening in the world. You know, flowers growing, leaves fluttering or all sorts of interesting things that we observe. *Galileo* presented what were called experiments, but many of them were just invented thought experiments. It's now known that he didn't - in fact couldn't - have carried many of them out. And in fact, the famous experiments almost certainly were never even tried, like the dropping two balls off the tower of Pisa to show that different masses would fall at the same rate - it's unlikely that he carried it out. If he had carried it out, it probably wouldn't have worked. There's too much complexity in the world. If you actually look at his notebooks, what he did was to give a purely conceptual analysis. He said, suppose you had two balls of the same size and mass and you dropped them. Well, obviously they're going to fall at the same rate. And suppose you bring them a little closer together. Well, that's not going to change anything, so they'll still fall at the same rate. Suppose you bring them so close together that they're touching in one point. Ok, well, that shouldn't change anything, but now you have one ball of twice the mass, so it follows that mass isn't going to affect rate of fall. And much of the experimental work was like that.

It was hard to convince people that they ought to be puzzled about things that looked simple and obvious. And some of the things that seemed simple and obvious *Galileo* could never really explain. So for example, if the earth is rotating, why don't objects fly off into space? It's not easy to understand that. Later, much later, an explanation came forth. But in fact just about everything you look at that's happening in the world is puzzling. If you're willing to be puzzled by simple things, then that's the way science starts.

Well, it's the same in studying what we now call the Cognitive Sciences, human mental faculties. And so, for example, when *Descartes,* which really opens the

modern period of Cognitive Science, shortly after Galileo, and part of the same intellectual movement, asked - he imagined - an experiment, which, I don't know if anyone's ever tried out, but it would almost certainly come out the way he suggested. He said, suppose that you take an infant who's never had any experience with geometrical figures and you present the infant with a triangle, so – *[crowd laughing]* use your imagination, suppose - it's all imaginary anyway - so like the balls on tower of Pisa, so, *it's ok don't worry*, suppose I were to draw a triangle on the blackboard. Well if you look at it carefully, it certainly wouldn't be a triangle. The lines would be bent and they wouldn't come together in the right place and so on, but *Descartes* simply says, if you present an infant with a triangle, the infant will see it as a distorted triangle not as a perfect example of what it is. What it is, is some complex and indescribable geometrical figure but the child sees it as a distorted triangle. So *Descartes* was giving an answer to a certain question - which is puzzling - why does the infant, who's never had any experience with geometrical figures, see it as a triangle? He was answering what you might call the *what* question about the visual system. It's a question of what it is. Here's the way it works. The way it works is that the system sees things as geometrical figures, maybe distorted geometrical figures. And then he asked another question. Why is that the case? And he gave an answer, which probably isn't the right one, but made some sense. He said, well, the child has an innate, we would say genetically determined, structure in the mind, that determines the way the visual system works and it essentially incorporates the principles of Euclidean geometry. So, that's the only way to interpret things, as distorted Euclidean figures.

So that's an answer to the question *why* it's that way. And then you might go on, now being kind of anachronistic, and ask a question how that evolved? And there you get into quite interesting questions. Surely it didn't evolve through natural selection. It evolved because of some way physical laws operate in the

development of complex systems, like the visual system. At this point, you're reaching the questions that are right at the fringe of contemporary evolutionary theory. How do physical laws enter into determining the channels through which evolution can take place and the nature of the system?

The same attitude of being puzzled by obvious things and trying to answer questions about *what* the system is, *how* it develops in the individual, *why* it works this way and not some other way, now, those are very live questions on very simple issues, like this one for example. Well, that was the point of view that. it seemed to us, one ought to take toward language. It became known later as the *biolinguistic* framework, but doesn't need a fancy term. It just says, let's look at language in the way any scientist would look at any biological system. Language is some biological property of human beings. As far as we know, a species property, meaning identical across the species apart from very marginal fringes which we call pathology. But essentially identical across the species and unique to the species. We don't know of anything comparable in any other species. So it's somehow a defining characteristic of humans and we should study it like the way we study the visual system and any other system.

That means asking questions about what it is, how it develops in the child and why it evolved that way and not some other way. And where do genetically determined elements enter into it, as distinct from others. Well, that yields what's called the problem of poverty of stimulus. The poverty of stimulus just means that there is a gap, a huge gap in fact, between the data available to the child, and any other organism, and the competence that comes out, what's known at the end. So in Descartes' case, for the specific question he posed the gap is a hundred percent. He imagined an infant who has never seen a geometrical figure, so no data. And what comes out, according to his reasoning, is a distorted triangle, not the actual figure that's there. And the answer to why it

works that way would be that there is a genetic element that incorporates Euclidean geometry. In modern terms, we might want to, say, speculate now, because no one really knows to the extent that that's true or whatever the right answer is, probably not that, it's because that's the way the laws of physics work. There is no other way for a visual system to develop. Those are possible answers, maybe wrong, but those are the kinds of questions that have to be asked and the kinds of answers that have to be given.

And the same is true about every aspect of human language that you look at. Poverty of stimulus problems are ubiquitous. Every aspect of growth and development poses huge poverty of stimulus problems. Now the term isn't used in biology and the reason is it's taken to be so obvious that there is no need for a term, so it's obvious that there is a poverty of stimulus problem when humans develop arms instead of wings or a mammalian visual system but not an insect visual system. There is a stimulus. There's external data like nutrition but there's – no one even bothers to argue about it - there is no way for nutrition to determine that you have a mammalian visual system so that's got to be accounted for by something internal, some genetic property. And then you go on to try to find out what it is and ask why it's that way and not some other way. In the case of language, there is a term, poverty of stimulus, and it's considered highly controversial, but just about everything about language is considered highly controversial, even if it is perfectly obvious, a total truism. It's just a fact about the irrationality with which humans approach their own capacities. It holds across the board. Everything about ourselves seems obvious, you know, after all we do it, what else could happen? It's like the apple falling from the tree. And since everything seems obvious, how can there be a problem? So, when anything is suggested as a puzzle or a problem, that's considered trivial.

By the time, back in the 1950s, when this work was taking off, the standard view in linguistics and psychology and engineering and related fields was that it *is* trivial. That is, there are simple ways of answering all these questions. Through radical behaviorist concepts in the study of behavior generally; for language, structuralist procedures, which automatically give you the results when you apply them to data. Information theory was then becoming a prevalent idea - in particular, not so much *Shannon's* actual work as Warren Weaver's supplement to *Shannon's* famous paper, which everyone read, which made it seem as if great discoveries were in the offing as soon as we applied these concepts. Computers were just coming along. There was tremendous euphoria about how, any day now, computers were just going to do everything. You just feed data into them, do some statistical analysis and out comes everything you'd want to know. There was like a six months' gap - we used to make fun of it - in six months, we will do x, y, and z, and the computer will be the same as the brain. So there was a kind of sense of euphoria, which added to the general sense among people that there can't be anything mysterious about what we're doing, there can't be any far-reaching puzzles.

But, when you look, it turns out everything is a puzzle. Many examples were shown, were exhibited, and many of them are still puzzles after many years of work, just as for example, the problem that troubled *Galileo* as to why things don't fall off the earth remained a puzzle for a long time. These puzzles are what I mean by unfinished business, aspects of some of the earliest most trivial examples, which are highly problematic, although the problem hasn't been recognized.

It's hard to notice puzzling aspects of things which are familiar and intuitively obvious, but it's worth bearing in mind that that's the way serious enquiry proceeds. It begins by being willing to be puzzled about things that sort of seem

natural and to ask well, why is it that way and not some other way, as in *Descartes'* case, or in *Galileo's* cases or contemporary ones.

Just to illustrate, I'll take one example that was presented back in the 1950s and has become a sort of a classic case, because it's so trivial. It seems that we ought to understand it perfectly well, and in fact nobody had ever noticed that it was a puzzle before and many still don't consider it a puzzle, but I think it is. So take a very short sentence, take the sentence *[Chomsky is writing on the blackboard]* "*Can eagles that fly swim?*" Ok, simple sentence. Everyone understands it. Any young child understands it. There is a question about it. We know that we associate the word '*can*' with '*swim*' not with '*fly*'. We're asking "*Can they swim?*" we're not asking "*Can they fly?*". Well why is that? A natural answer ought to be that you associate '*can*' with '*fly*'. After all, '*fly*' is the word that's closest to '*can*' so why don't you just take the closest word and interpret it that way? Well, you obviously don't. You interpret it with '*swim*'. So to illustrate, just as a notation for now but it turns out that is not just a notation, there could be a - I'll put a star meaning it doesn't happen [*writing something on the blackboard*]. You could interpret it in this position as a verbal element, a verbal auxiliary, now related to '*fly*'. But you don't. You interpret it in this position as a verbal element related to '*swim*'. Well, that property is universal. It holds up in every language. Languages may do it differently but they're going to have the same property. It holds in every construction anyone knows and it's just a universal property of language. Well, this particular example has taken on a life of its own. For one thing, it's a poverty of stimulus problem, like *Descartes'* triangle. There's been a huge effort to show that it's not a problem, that if you just do a complex statistical analysis of complex data, you'll find that that's what the child will determine from the data. The approaches are odd in several respects. First, every one is not only a failure but a colossal failure. I'm not going to talk about that. I actually have a recent paper about it with a computer

scientist at MIT, Bob Berwick, where we run through a lot of the current proposals but it's easy to show that they're all just wildly wrong. But they keep coming. Almost every issue of the main cognitive science journals has another proposal of this sort, so that's one odd fact. There are many efforts to show that there is nothing puzzling about it, they're all colossal failures.

The second odd thing about it is that there's a trivial answer. The trivial answer is that the relationships between positions are what are called structure dependent, not linear. Ok, so that if you give a structural description of this, pretty much like traditional grammar, a word '*can*' sticks out over here, and then there is a sentence *[Chomsky is writing on the blackboard]* and then there is a nominal phrase "*eagles that fly*" and then there is a verbal phrase '*swim*', and the '*swim*' if you actually look at it, is a little more complex. It has an inflectional element which is usually called *T*, to stand for tense, and then a verb phrase element. So if you had the '*can*' and just a declarative sentence, it would be "*Eagles that fly can [Chomsky shows and writes on the black board*] - that would be here – '*swim*' would be here, so that's sort of straight out of elementary school grammar. In contemporary generative grammar and related areas it is called phrase structure grammar. But it basically draws from tradition. Now if you look at that characterization, then there's a simple reason why you should get this association *[Chomsky showing something on the blackboard]*. These two positions, where the word '*can*' could appear, happen to be very closely related structurally, in fact, the closest related structurally, the relation between '*can*' and what I wrote as "V star" is a much longer relationship because it's embedded inside the phrase. And if language uses structural closeness rather than linear closeness, then this would be the answer.

Well, this is kind of parallel to the Descartes' argument. First question is what happens? What happens in this case is, you interpret the auxiliary as being

connected with position of the main verb not with the embedded verb, that's what happens. How does it happen? Well the child uses principles which were assumed to be genetically determined, which require structural closeness rather than linear closeness, so that's like the child having Euclidian geometry as what we would call a genetic property. But then comes the *why* question, why does it work this way? Well, that's not so obvious. One aspect of why it works this way *has* been discussed. But there's another aspect of why it works this way which hasn't even been noticed. It's another one of those cases where things which seemed too simple don't elicit the puzzlement that they should. The first question has to do why you use structural distance not linear distance? Well, a possible answer to that would be that for principled reasons which we'd have to figure out, languages just don't have counters. They don't use linear properties at all. Then that raises another *why* question if it's true: *why*? Well the answer could be, and probably is, that the semantic system of language doesn't have order, doesn't have order at all. In the system of thought that you use, there is just hierarchy and structure but no order, no temporal order. That's pretty natural if it's true. It's hard to show that it's true. But if it's true, it would mean that linear order derives from the fact that order is a property of the sensory motor system. One elementary truism about language, which like a lot of truisms turns out not to be quite true, is Aristotle's description of language as sound [*is writing something on the board*] with meaning. There's something right about that. So language has sounds and has thoughts associated with it. I'll raise a question whether it's true later but at least it seems to be true. It's always been assumed to be true. Now the sound side of language involves the sensory motor system, and the structure of our sensory motor system is such that we can't speak in parallel. We just speak linearly. Actually, there is a kind of dolphin, which I've always envied - dolphins blow air out of their nasal passages - there is one species that blows them out of both sides at the same time and it's used for its communication system, so essentially it can speak out of both sides of its

mouth, more or less. And it can speak in parallel, so ok, that wouldn't have to have the kind of linear order that we have. We're not that well constructed.

Sign language, which is just like spoken language, it's now well-understood, does have something similar. There are parallel actions taking place in sign language. So if you raise your eyebrows, let's say, you can raise it over a whole clause and that tells you that the clause is a question, not a declarative and so on. So that does have more options. But spoken language is linear, so it's pretty natural to assume that the linear aspect of language is just kind of a reflex of some other system. The fact that it's coming out spoken language, using the sensory motor system, doesn't allow anything else. Well if that's true, you wouldn't find it in the system of thought and you wouldn't find it in the system of principles that determine the structures that enter into thought, like this one. That's very hard to show - it's natural - but very hard to show. There is a lot of empirical evidence against it. And what the task of the scientist would be at this point is to show that the empirical evidence is misinterpreted. Since this ought to be true, let's assume that it is true, and then try to see why the apparent counter evidence is wrong. That's the way almost all of science proceeds. Almost anything you do at first comes out wrong and what you try to show is that where it appears to be coming out wrong, you did the experiment incorrectly, you're looking at the wrong data, something's wrong. That's the way rational enquiry proceeds and it's how it ought to proceed here.

Well, that's one part of puzzle that still has to be answered and it hasn't been. There are current papers, technical papers, that seem to show that linearity is involved in the construction of expressions, though not apparently in the thought system. That's initially implausible, because there's no reason why it should be in the construction of expressions if the thought system doesn't need it, merely because the sensory motor system is requiring it. So there is motivation to try to

show it's wrong. Task for the future. However, there is another puzzle here which has not been noticed yet, until recently, and that is that there is something arbitrary about that structure. That structure tells us that the most prominent element is '*can*', what I called *T* and that's what you get out of traditional grammar and that's what modern phrase structure grammar assumes, and that what's contemporary papers assume, but there is a stipulation there. Why, for example, doesn't it look like this? [*Writing on board*] Then the head of the noun phrase, let us say a determiner *D*, is the structurally closest element to the pre-sentence position. With that structure, the yes-or-no question asking whether *many eagles that fly swim* would be *eagles many that fly swim*. One might argue that that should be expected, given well known similarities between clausal and noun phrase structure, but it is obviously wrong. The structurally most prominent element is *T*, not *D* – and if it happens to be empty, the semantically vacuous element *do* is inserted and the corresponding question is *do eagles that fly swim*. But that raises a puzzle. Why don't we do it the other way?

In technical terminology, the noun phrase is called a specifier of *T*. But the question is why don't we call the TP the specifier of *D*? It could be either way, so why do we it one way and not the other way? Well, that's the way that makes sense but that can't be the answer. That's like saying an apple falls to the ground because it makes sense for it to do so. So there's got to be some answer to this. And it turns out that it is not a trivial problem and when you look at it you get in all sorts of difficulties.

I'll get back to it and talk about what the difficulties are and the consequences. But first, let's look at the whole problem a little bit more systematically. So let's take a look at Aristotle again. Language is sound with meaning. It's a good starting point.

In fact practically the entire history of the study of language for the last 2500 years has been the study of sound in a broad sense: phonetics, phonology, inflection, morphology, linear order, and so on, all having to do with the way things come out the mouth. There is a little bit of study of meaning but I think it's mostly off track. It's almost entirely based on the fundamental assumption that there is a relation between words and extra-mental objects. So that if I refer to a person, there is a relation between my name of that person, an internal symbol, and a physical object that a physicist could identify. But it just takes a moment of reflection to show that that's not true. That's not the way words work. That's not the way meanings work. And furthermore every infant knows it. So a standard infant's fairytale is that the prince is turned into a frog by the evil witch and every single physical property of the prince is that of a frog until the beautiful princess shows up and kisses the frog and all of a sudden the prince comes back. But what the child knows is it was the prince all along. It wasn't a frog, it just had all the physical properties of a frog. It was the prince all along and that's why when the princess kisses the frog, it becomes a prince not a horse or something. That's something that's so elementary that every infant understands it. But if you think what that means – that means that we pick out persons by a property that has no physical characterization, namely psychic continuity. We pick out persons by attributing to them psychic continuity, and we do the same with animals. It's easy to show parallel situations and children understand it instantly. For example, there is a children's story that my grandchildren like. It's about donkeys. There is a baby donkey who's turned into a rock. And it's a rock by every property you can imagine. It spends the whole story trying to convince its parents that it's their baby donkey. But they don't understand and finally - children stories have happy endings – that's an innate property – and so somehow it turns back into the baby donkey and everybody is happy. But the child who is hearing the story understands that that rock is a donkey. It's not a rock. Even though every physical characteristic is that of a

rock. And in fact every word of language that you look at is that way. You may think that when you use the words Rhine River, you are talking about a physically identifiable object but it's quite easy to show that you're not. The properties that determine that it's a river not a highway are extremely subtle and are not physically identifiable, and so on with just about every word you think of.

So the whole study of meaning is based a serious fallacy and the facts pose enormous poverty of stimulus problems. Like how does a child know that psychic continuity is the way you individuate persons? Where is the evidence for that? This is a huge poverty of stimulus problem. And it also raises a problem for evolutionary theory, which hasn't been addressed. Maybe it's beyond investigation but this seems to be a specific property of human language, which doesn't show up in the animal world – anywhere. There's been a lot of study of animal communication systems and they really do have this referentialist property. The signal is one-to-one correlated with some physically identifiable extra-mental entity. A vervet monkey will give a certain cry if the leaves are moving in a certain way – a warning cry. And it does it reflexively – if the leaves move, it gives the cry. And that's identifiable. A chimpanzee gives the signal that says "I'm hungry" if it's hungry and can't avoid doing it but there is an identifiable physical state. Humans are nothing like that. First of all, our concepts don't relate to the world that way and of course they are not produced reflexively. You can look at the Rhine and say it's a mountain if you want to or anything else or say nothing. So there is some special of property of humans, right at the core of the conceptual system, that has no relation to anything in the animal world and that's of totally mysterious origin. No one even knows where to look for an answer or if it's possible to find an answer. But that's again a major poverty of stimulus question that arises as soon as you begin to allow yourself to be puzzled.

Well let's go back to the notion "sound with a meaning," and now let's try to take a systematic approach to this. While there has been a lot of study of sound and inflection and morphology and so on, and very little about meaning, there has been almost *no* study of the word "with". How do the two get connected? And until about the mid 20th century, it wasn't very clear how to address that question. One elementary property of language is that it is a system of "discrete infinity". Sentences can be arbitrarily long but they are like numbers. You have natural numbers – one, two, three, four – but three and a half is out of the natural number system. And it's similar with sentences. You could have a five word sentence and a six word sentence; you can't have a five and half word sentence, and goes on indefinitely, like the number system. So it's a system of discrete infinity, which is very rare in the biological world, in fact, it may be unique. There's nothing analogous to it. Well, it has been understood well for about 50 years how to study systems of discrete infinity. There must be what's called a recursive procedure, a generative procedure, which forms an infinite number of structured expressions (or some equivalent), and the mathematical theory of such procedures is quite well understood. Well if you look at any such procedure, you're going to find buried somewhere in it an operation which says "take two objects that are already formed and make up a new object from them". So take x and y, which have already been formed, and apply some operation to them, which yields some new object z. Somewhere in every generative procedure you are going to have that operation. So that's going to be in language somewhere. We would like to show that the operation is as simple as possible, both for methodological reasons and for pretty good evolutionary reasons, which I'll come back to. We want to show that it is as simple as possible. Well it's going to be as simple as possible if you don't change x and y in the process. So that means that z is, in fact, nothing but the set containing x and y unchanged. That's the simplest computational procedure so, unless we have evidence to the contrary, we'll assume that's the core procedure of natural language. Well that

has a name. It is called "Merge". So Merge of *x* and *y* gives the set {*x*, *y*}. Simplest possible operation, so short of counter evidence, that's what we assume. That's the first step towards explaining "with".

Just by pure logic, there are two kinds of Merge: *x* and *y* can be distinct from one another or they can be not distinct from one another. If you look at the process by which Merge operates, the second case where they are not distinct means one of them is inside the other – no other possibility, just logic again. So those are the two possibilities – either *x* and *y* are distinct or one of them, lets say *x*, is inside *y*. There are names for those operations too. "External Merge" is the case where *x* and *y* are separate, so if you take, lets say, "eat" and "apples" and you externally merge them, you get the set {*eat*, *apples*}. That's one possibility. Suppose one of them is inside the other. So suppose, for example, you're trying to form the expression "what John ate", as in "Guess what John ate." Well, you start with "ate what", you externally merge to it "John", giving what underlies "John ate what," and then you take "what", which is inside this, and put it over here and you get "What John ate what". That's what happens if you apply internal Merge without modifying anything, the simplest possible way. Now that gives what's called the "copy theory of movement". You end up with two copies. And for semantics, that's exactly what you want for the sentence "What John ate". What you want for the semantic interpretation is "What John ate what" because you have to know two things about *what*: you have to know that it's an operator that ranges over the whole sentence but also that it's the object of the verb "eat" and that's where it gets its semantic interpretation. So it has to appear in both places and that's what you get automatically from just applying Merge in the optimal fashion.

Well, there's a consequence here: you don't say that. Nobody says "What John ate what" and the same holds universally. When there is something that has two positions of interpretation, you only interpret it once. The same in the "can

eagles that fly swim" case. You understand "can" in two different positions. The initial position which says 'I am asking a question' and the position next to fly which says 'I am asking about fly not swim'. Well, that's internal Merge again. You have just applied the most elementary existing rule in the most elementary fashion. And you get the two positions which is just what you need for semantic interpretation. But you don't pronounce it.

At this point, sound and meaning begin to separate from one another. The simplest design of the system based on Aristotle's "with" happens to work fine for meaning but very badly for sound because you don't pronounce it that way. There's an asymmetry between the sound side and the meaning side, which is crucial. Now actually pronouncing just one occurrence is also highly economical. If you look at what goes on in the brain, let's say, when you articulate something, there is an awful lot of energy and activity involved and you can minimize that by just not doing it. So the easiest way to use a language is just to pronounce it once. Well, you have to pronounce it at least once or the hearer has no evidence that you formed a question or some other operation. So you have to pronounce it at least once and you also have to pronounce it in the hierarchically most prominent position: otherwise you'll get the wrong interpretation. You'll interpret it down below somewhere. And that's in fact the way language universally works. Every language universally keeps the copies there for interpretation and in fact if you get non-trivial sentences, turns out that there is many copies scattered all the way through and it uses all of them for interpretation but only pronounces one of them, the hierarchically most prominent one.

There are some interesting exceptions that provide further evidence for the conclusion that only the structurally most prominent copy is pronounced unless some extrinsic factor requires a residue of others to be pronounced.

Well, that's very good from the point of view of computational efficiency but it's very bad from the point of view of communication. The person who hears one of these sentences has to figure out where the gap is. In this simple case it happens to be easy but if it's is a complicated sentence, it wouldn't be easy. Any of you who work on parsing programs, trying to figure out methods for mechanical interpretation of sentences, knows the big problem is what is called a filler-gap problem. You hear the word "what" in the beginning of a sentence and you've got to figure out where the gap is. Well, that's also a perceptual problem, the problem that any perceiver faces. So to put it simply, what turns out to be the case universally is that language is very well designed for thought and very badly designed for communication.

Now there are many cases where there are conflicts between communicative efficiency and computational efficiency and in every case that's known, computational efficiency wins hands down. That tells you something about language; first, it tells you Aristotle is not quite right. Language is expressions with meaning, and sound is sort of tacked on there somewhere on the side and it doesn't work very well. So there's a fundamental asymmetry. Well, I won't give other examples but if you think it through you'll find that it's universal. Every language is like that. It's automatic.

Well, it might sound surprising, but going back to the *why*-questions, the evolutionary questions, it's very natural. It's what you ought to expect. So you can't say much about evolution of language, almost nothing. I think that's the reason why there are so many conferences and volumes and libraries and so on, there's essentially nothing to say – so you have a lot of talk about it. But there are a few things that you can say and the few things happen to lead just to this conclusion. So we know something about the evolution of language. If anything,

we know it was finished before humans left Africa. All human groups are essentially identical with respect to cognitive faculties altogether, but the language faculty in particular. So sometime before humans left Africa, it was all over. That's about 50 000 years ago - you can pretty well time that. So before 50 000 years ago, it was already done from the point of view of evolution. You can also pretty well time – this is more speculative - the earliest part. There happens to be evidence of rich cognitive symbolic activity in the archaeological record, roughly 75 000 years ago, maybe 25 000 years earlier, but sometime roughly around then. It's generally assumed by paleoanthropologists, and plausibly, that this must be because language developed. So roughly at that point, you get the appearance of complex symbolic behavior, symbolic art, complex social arrangements, recording of natural phenomenon and all kinds of things. It's sometimes called "the great leap forward". Roughly at that period, all of a sudden (in evolutionary time), something exciting happened. Well, that pretty strongly suggests that language developed around maybe between 100 000 years ago and 50 000 years ago. Now, you can make the stretch longer if you like, it doesn't change much. The point is from an evolutionary point of view, this is virtually instantaneous. There is no time for selectional processes to have had much of an effect because it's just too quick. So what must have happened - must is a strong word - but the only plausible assumption is that sometime in that period, suddenly the concept Merge developed. Once it develops, you have the capacity to form complex structures, structured expressions, indefinitely long, and they can then be related to the conceptual apparatus, which obviously existed even though it's a complete mystery why it has the properties it does, as I mentioned; but it plainly existed. So you could link them up and you could think. Well, that means that some mutation took place, that's the way changes take place, so some mutation took place, maybe a small mutation which gave that capacity. Well, mutations take place in an *individual* not in a group. That means some individual was lucky enough, or maybe unlucky enough, to

suddenly have the property Merge. That individual could think. It could form complex thoughts. It could plan. It could interpret and so on. That gives some selectional advantage. Mutations that have a selectional advantage *can* propagate. It's not simple. Most of the time they don't, but they can propagate obviously, that's why evolution takes place. And they propagate through the descendents, so some descendants had the same property. After a while maybe some series of generations, some of them might have had the bright idea to try to externalize what's going on in their heads. There is no point in doing it if you are the only person who has this capacity, then nobody would understand it if you did. But if there are enough people who have that capacity, then you can usefully externalize it. Well, that story looks almost exactly like what we discover. Namely, the core systems of syntax and semantics, constructing expressions and interpreting them, they seem close to optimal. The more we come to understand, the more they seem to be optimal. That's exactly what you'd expect on this scenario. Once Merge appears, there are no external pressures at all determining how it develops. It's just sitting there. It's going to develop just by the laws of nature. One of the laws of nature is computational efficiency, so it will develop through computational efficiency and therefore it ought to be essentially perfect. What's called the minimalist program is just an effort to try to show that what ought to be true in fact is true, or close to it. There are plenty of impediments, too numerous and difficult to mention. But there is progress. And it's a reasonable research endeavour, because something like that is what you'd expect.

But what about the relation between this internal system and the sensory motor system? That's the externalization problem. Well, the sensory motor system had been around for hundreds of thousands of years. It's a completely separate system. It has nothing to do with this internal entity. So there is a hard problem to solve. How do I relate that internal system to the sensory motor system for

externalization? Well, it's a hard problem and in fact if you look at language, that's where practically all the complexity of language is. When you study a second language, about all you study is externalization. You study the sounds, the particular lexical choices, which are arbitrary, the inflectional system, you know, how to conjugate verbs, some facts about word order, and so on. That's just about all you have to learn. You don't have to learn the syntax and the semantics because that's there already. That's part of your nature and probably it's part of your nature because that's the way physical laws work. It's meeting conditions of computational efficiency – or so we would like to show.

The externalization systems are overwhelmingly – maybe, some day, we will discover *entirely* -- where languages differ from one another. The wide variety of languages is almost entirely, maybe entirely if we know enough, in the externalization process, the secondary process of getting it out into the sensory motor system. That's also where languages are very susceptible to change, so say teenage jargon or invasion or something else. That's where languages change a lot. That's where they vary. That's what you have to learn because it's pretty much arbitrary – not totally. You probably solve the cognitive problem as simply as possible but it is a hard problem, a problem that every infant has to solve. I mean an infant is getting a lot of data. It has to connect it to this, probably fixed, internal system which is determined genetically but probably by virtue of laws of nature to a large extent - not entirely, because at least Merge must be part of the genetic component, what's called universal grammar, and maybe other things, but it shouldn't be too much. The less that's in there, the easier it is to solve the evolutionary problem and since it's almost instantaneous, it must be easy to solve if we ever learn enough.

So that's the general picture. Though there are huge gaps, it kind of hangs together. It means that language is not an instrument of communication.

Contrary to what is universally assumed, it's not well designed for communication. But it's well designed – maybe even perfect -- for expressing thought. There are principled reasons to believe something like that and it seems to conform well with the facts that we know. It means that almost the entire study of language for 2500 years is kind of off track. It's studying a secondary problem, namely how the sensory motor system links to an internal system that is language. That's not, strictly speaking, a linguistic problem. It's a cognitive problem in substantial measure, do it appears.

Let's get back to the original puzzle that I brought up earlier. If the picture I just outlined is on track, it would suggest an account for one of the two puzzles about "can eagles that fly swim": why you use structural rather than the linear minimal distance. If the general picture is correct, there can't be a linear property. The linear property all has to do with externalization and that's where languages differ. German puts the verb here, English puts it there and so on. It's the kind of thing you have to learn when you learn a second language but there is no evidence that any of that enters into the thought system. Sentences are understood exactly the same way if you put the verb at the end or at the beginning or in the middle and so on. So that's probably just one of the solutions to the externalization problem. If that's correct, and it's far from trivial to show, but if you can show that, it ought to be correct, so that means you should be able to show it with enough work, that will account for one of the puzzles. Why do language use structural distance rather the linear distance? The reason you use *closest* is probably a law of nature. Laws of computational complexity say "do things as simply as possible" and closeness is as simple as possible, there's less search. So that deals with half the problem, but now we have the question of why we regard the *D* head of the noun phrase as hierarchically lower than the *T* head of the Tense/verbal phrase. Why don't we regard the verbal phrase including T as hierarchically lower than the head of the noun phrase, say the

determiner? Why that asymmetry? Well, the simplest answer to that is that the subject of the clause isn't there when the relationship is established.

There is good reason to suppose that extra-clausal position where *can* lands is already there. Technically it's called the complementizer position. The complementizer C tells you what kind of a sentence you have - do I have a declarative or interrogative, imperative or something else? That's the C position. The C position and the TP must be there before the subject is introduced. In that case, the asymmetry is accounted for. The C and the T will indeed be the closest related things because there is nothing else there, so that would solve that puzzle. Now it turns out that there is a very broad assumption in the field that the subject is actually inside the verb phrase and it moves to the position adjacent to the verbal phrase. It is called the predicate-internal subject hypothesis. And now we have an argument that says that must be true. In fact, that's the first real argument that shows that this must be true. There has been some evidence for it and it's been almost universally assumed by now because it's plausible on other grounds, but there was never really any good argument for it. So we've now found solid evidence for that property with a lot of good consequences that follow. But there is still something wrong. This analysis assumes that in forming the sentence, we first have the C and then we have the TP, which includes the verbal phrase which contains subject, and the subject has to move to this position between C and T. Well, that's an operation that's called countercyclic. It doesn't just add things to the edge; it puts things in the middle. There's very good evidence that language doesn't work like that, nor does cognition. Constructing new things, you stick them at the edge of what you had. Otherwise it's a kind of tampering, a more complex operation. So how could that be true? Well, I don't have time to get into it but there is a pretty plausible theory called phase theory, which predicts that that's exactly what ought to happen. That in precisely this case, the operation looks countercyclic, but in fact is not. The subject is trying to get to the C position but it's forced to stop at the T position

because that's where agreement takes place and so on. So that theory is sitting around, there is good evidence for it. Now we have better evidence, it's forced to be true.

Well, I think this goes on a long way. Once you eliminate the notion specifier from language, the notion of these asymmetries that are taken over from traditional grammar but in fact are stipulations, it turns out that all sorts of things change. You have totally new problems, interesting solutions engendering more new problems. We may be able to take care of some unfinished business, but typically that reveals new and unsuspectied unfinished business, the process that makes science perennially interesting. But I think I have talked enough so I'll stop there.

Q: Do you think that externalization of language also has to do with genetic mutations or do you think that it just uses resources that are already there?

N: We can really only speculate. Almost everything about evolution is speculation. I mean, you know everybody assumes that Darwin's sort of basically right but if you try to demonstrate the effect of natural selection it's extremely hard. There are very few cases, any biologist will tell you, where you can get real convincing evidence of selection. There are cases of plausibility. So now we are talking about a plausibility question. So let's imagine this small community, because the communities were pretty small, of hunter gatherers somewhere, a scattering of whom have this internal capacity to think, you know? How do they get to the point of deciding to externalize it through the sensory motor system? Well, it could have been another mutation or it could have just been what comes to mind when you have a capacity and you see – you have evidence that other people see it, they are also planning and interpreting and so on – so you figure, well, they must be like me. Our usual assumption about other people is "They are like me". That's called theory of mind. That's how you understand other people, sympathize with them and so on. You assume "they are identical with me". So if they are doing, something it must be because I would have done it under the same circumstances, and I attribute to them the same properties that I have, and that's the way we more or less get along with each other - and not too well sometimes but to the extent that we do. So it's quite plausible to assume I think that in this little community, if a few people could see that someone else is interpreting and planning the way they do, the way I do, that they are probably the same as me. Just as we interpret anger, sympathy, fear, or just about everything. And if they did, then it could easily come to mind that I ought to figure out a way to make public what's in my head. That means to map it into the sensory motor system. And then how do I do that? It turns out that it is not so simple because, as I said, the sensory motor system has been there for

hundreds of thousands of years, we have good fossil evidence for that, totally independent of whatever was going on in the head. So I have to solve a problem. I have to solve the problem of how to externalize this. And then once I have done it, other people are going to have to solve the problem of how that noise out there or sign or whatever it turns out to be, relates to the internal system that they have. That's the problem every child faces today. Every child faces the problem of figuring out how that data out there relates to this internal system that I have, which is probably uniform. It's probably just a fixed species property. And that's language acquisition. It's a fact that it takes place almost reflexively so there has got to be some simple solutions to it. But it is subject to variation. Like if you grow up in Mainz, and you grow up in Cambridge, it's going to turn out differently. So different solutions to the same problem. That's very much like other biological systems. We know experimentally that we all end up with pretty much the same visual system but that's because we have pretty much the same data. But if you manipulate the data experimentally, you get radical changes in the structure of the visual system. It's been shown experimentally for cats and monkeys. In fact, that you get total destruction in the visual system. So for example, if a kitten doesn't see patterned stimulation, they can have plenty of stimulation, but if they don't get patterned stimulation in the first couple of weeks of life, then the peripheral parts of the visual system actually degenerate. They don't function anymore. There is no way to ever to see. That's a radical change. But you can get less radical changes just by modifying the visual environment and that seems to be what happens with language. The environment varies but that's probably because there are just different ways of solving this externalization problem. And then each child comes along and has to try to make some sense out of it, and it changes a lot; you get quick historical change and so on. One of the interesting things about language acquisition, which nobody really understands, is that children tend to talk like their peers not like their parents. So it's not that their parents are

instructing them. If you think about yourself, you probably talk like the kids in the neighborhood and not like your parents. And somehow that just happens. That's just an automatic thing about the way humans develop. We're now well beyond what anybody understands but it's overwhelmingly the case and it's just the way these cognitive processes function. A lot of puzzles there. A lot of good questions to try to get PhD theses on but it happens and it really hasn't been studied as far as I am aware. As far as I know, it has never really been studied but it happens uniformly. So that's just a speculation, could have been another genetic mutation but that's a little hard to explain because that mutation would have had to take place in everybody. And what would cause that? What would cause the mutation to take place in every one of these scattered number of individuals who entered into the externalization game. It doesn't look likely. It looks more like normal empathy, which we don't understand, but we know it takes place. That's the way we interpret what other people are doing. We basically assume they are identical to me, and if they are doing something, it's probably what I would have done in those circumstances. So what are the circumstances? Then you attribute it to them an internal state like the state that would've led you to do it. That's so normal that this could well be a case of it.

AF: But, even when taking away the externalization aspect, would that then mean that even our closest relatives would not have Merge?

N: It would mean – first of all the closest relatives are not so close; like, let's say, chimpanzees are like 12 million years or so, remember it's not seven, it's usually said six million but that's separate evolution both ways, it was like 12 million years which is not huge but substantial, and it gets even worse when we think that in our own species, the evolution is apparently very recent long after the separation from existing primates, so millions of years after the separation from existing primates, this seems to have happened. So there is no reason to

expect anything similar in other primates. Which means that all the work that's done on trying to torture poor monkeys into having language makes about as much sense as trying to train humans to fly. It makes even less sense because at least we have the physical equipment to fly, like you can imagine moving your arms in such a way they could make you fly. But there is no reason to think that they even have the physical equipment that would enable them to do it. So it's not really surprising that everything fails completely. There are just no analogues. They don't even seem to be analogues at the most elementary level, like concepts, like the concept river or tree or table or pen or a person or anything else. Even there, there don't seem to be any analogues in the rest of the animal world. That's a mystery but its just overwhelmingly true. And no children have to be taught this. In fact, second language learners don't have to be taught it. You wouldn't know how to teach it if there were some need to do it. They just *know* it automatically. So it's something about the nature of the beast. There may have been other hominid groups that had the same capacity but there is another fact that's worth remembering about humans: we're a very predatory species. Anyone anywhere near us throughout evolutionary history got wiped out. It goes back to Metazoa, big animals, mammals and stuff. As soon as humans, proto-humans - this is long before Homo sapiens – came along, everything else got destroyed. So the chances are that if there were other species like us – relatives – they probably just got destroyed by this group that had a little better capacity.

Q: How can you demonstrate that those principles which you've demonstrated here for the English language actually apply to all languages? And how would you reply to your peers who try to find languages somewhere in Micronesia - in some remote part of the word - where those principles don't apply.

Q: If you basically say that language in principle is a means to think, would you then say it's the *only* means in which humans can think or would you say that there are other means, such as thinking visually, for example?

N: On the second question, it's really – it's not well formulated enough to have an answer. To answer it, we have to figure out what you mean by "think". We don't have any answer to that. It's just too broad and loose a concept. I mean language is a primary means of what we call thought but there are a lot of things we call thought and they may be heterogeneous and you'd have to clarify the concept of thought before you can try to answer. It's too loose.

On the first question: it's a little bit like asking "How do we know that all infants..." – lets go back to the Descartes. As I said, I doubt that anyone has actually carried out his experiment but I don't think that anyone doubts that it would come out exactly the way he thought it would. That's probably why nobody's bothered to carry it out. So how do you show that every child is like that? One approach would be to study every child but no scientist would do that. You just assume that every child is like that because we all the same cognitive capacities. And that's the standard approach of the sciences. To be more explicit, if these linguists can show that the principles don't apply, then either our formulation of the principles is wrong or else their conclusions about the languages they are studying are wrong. It's like asking chemists how can they demonstrate that their principles apply to hitherto undiscovered elements. They can't.

Take some remote tribe, let's say somewhere in the Amazon, which has never had any contact with others. Take a kid from that tribe and raise him in Mainz from infancy, it will talk exactly like the people who grew up here. It hasn't been tried with every child in the world but the evidence for that is just

overwhelming. And it's exactly what we'd expect because all these capacities were fixed before some small group of our ancestors left Africa and spread all over the world and all the properties were fixed by then. So you just don't find differences in cognitive capacities generally. In fact, you know, there is very little genetic variation among humans altogether. We are a very uniform species. We look for differences, like skin color, hair length and so on but the differences we look for are extremely superficial. They make a big difference in human life – but that's our own craziness. The basic properties of humans seem to be identical, close to identical, except for pathology, which you can find anywhere. So you could investigate every individual, just like you could investigate every apple to make sure it follows the laws of motion when it falls, but nobody does that because there is so much evidence that they have to be identical. In fact it's interesting that these questions only arise in the human sciences. They never arise in physical sciences. It's all the same. Why don't they arise elsewhere? Because there is a kind of rationality that prevails in the study of the natural world, which is somehow cancelled when we study ourselves. At that point we become very irrational. So we ask the kinds of question that wouldn't arise in studying other aspects of the physical world - even studying other animals. Like, take what I said about kittens' vision. I don't know actually how many kittens were studied, some small number, maybe 20 kittens. But it's assumed without studying any other mammals that that's the way all mammalian visual systems work. Hasn't been tried, even with kittens, but there is just no other plausible assumption. There is no counter evidence, and there is no reason to expect any. So the answer to the question is: it's a very common question that's raised, a very standard question, but we should think of it as a sign of the fundamental irrationality of the way we look at humans - different from everything else. It's maybe natural that we look at ourselves differently from other creatures but it's not good for rational enquiry.

Q: This question may reflect my poverty of mind. Learning language is a significant effort, it takes many years. [NC- It's reflexive.] What is the hard evidence that this significant effort is too poor a stimulus and we have to rely on the genetic - on a innate program in order to learn the language? Why is this stimulus too poor?

N: OK. Take the example that I gave, "Can eagles that fly swim". The chance that any child has heard a sentence of that structure is extremely slight. The idea that every child in every society has heard sentences like that is minuscule and yet everybody gets the same answer immediately. Well, that's basically enough and that's what we find everywhere. Furthermore, the idea that it takes effort to learn language is very questionable. I mean -- Angela can tell you a lot more about this than I can -- but as experimentation has gotten better, it's hard to experiment with infants, but as the techniques of experimentation have gotten better, it's been consistently discovered that what was thought to be a later stage of development, is actually there already, right at the beginning. As early as you can get. And in fact, it seems that language learning is virtually reflexive and it's not a trivial matter. I mean think about language acquisition, you've got this one-day-old child. There's all kind of things happening in the environment, how does the child know that some parts of what's going on in the environment are language-related? Nobody knows how to automate anything like that. I mean, no other primate, no other organism can do it. If the child has say, a pet kitten or chimpanzee or songbird or something, that other animal can get exactly the same data. But it can't even take the first step of deciding that some of it is language-related. That's a very hard step, but a child does it reflexively and instantaneously. And most of the rest of what happens seems to be about the same. Now, in fact, things are learned very early. In fact some of it is intrauterine. It has been shown that very young infants, like 2-days-olds, about the earliest you can experiment, they can distinguish between the language of

their mother and another language, both spoken by a bilingual woman whose voice they've never heard. Now, once this was discovered, a lot of experimental work went on and it's not any two languages, so the languages have to differ in certain ways and be the same in certain ways and it seems to be mostly in prosody and intonation, which in fact are learned quite early. And the rest of language learning, well, the part that looks to us complex, is mostly the kind that we do all our lives, like learning new words. You know, you go to graduate school, you learn a lot of new words. Ok, but that's the kind of trivial part of language learning. The basic structure seems to be fixed extremely early. So, how do you know that it's a poverty of stimulus problem? Well, how do you know that having a mammalian visual system is a poverty of stimulus problem? Maybe there is something about the nutrition that determines that if you gave an embryo and fetus different nutrition, maybe we'd get an insect visual system. I mean nobody asks that question because it's too outlandish, but do we know? No, you don't know, it just doesn't make any sense. And the same is true here. As I said, on this simple example, trivial example, it happens to have elicited a huge literature trying to show that, that it in fact is just a result of some analysis of data. But you can predict in advance it's going to fail. Because why should it be true for every language and every construction that you get essentially the same property. It's like saying maybe nutrition determines a mammalian visual system. You can't prove that it isn't true, but it just doesn't make any sense. And in fact it particularly doesn't make any sense because of the very simple answer. The very simple answer is along the lines that I described. Computational efficiency working in a very specific way. So you have a reasonable answer that goes to fairly deep principles. You have, you know, a conceivable possibility that somehow data made the difference, in which case you have to face the near miracle that it works everywhere, for every child, and every language. So, you know, while you can raise the question, it's not a serious question unless

somebody gives a suggestion as to how that could happen. And there have been suggestions in this case, but they all collapse very quickly.

Angela Friederici: Ok, time has moved on and I think I have one question over there. Yes, you. You in the blue shirt, raising your hand already.

Person: Thank you. You said, Professor Chomsky, that you thought the externalization process, like which internal complexity was externalized through the sensory motor system, you would posit that as having taken place about a hundred thousand years ago. My question simply, do see any relationship then between the rise of language and increasing brain size vis-à-vis higher primates has to do with the development of internal conceptual complexity?

Chomsy: That's a good question. There is some evidence that there was a sudden growth in brain size at roughly a hundred thousand years ago. The dates of these things are pretty tentative – there's a big variability - but roughly a hundred thousand years ago, there appears to be a sudden growth in brain size and it's been speculated, not implausibly, that somehow a mere consequence of the growth of brain size yielded this capacity. Well, nobody knows anything - anywhere near enough about the brain to spell out what this might mean. But it's conceivable. It's conceivable that, say, Merge is kind of an exaptation. Somehow complexity of the brain grew for whatever reasons. It just got bigger. Maybe more cognitive tasks or something and a side effect of the growth of the brain was you suddenly got this operation. It's not impossible, but there is really very little evidence about this. Remember, we're talking about soft tissue and you don't find that in the fossil record. What you find is hard things, so you can make some guesses from the shape of the skull and so on. But it's pretty speculative, and the real problem is that not enough is known about the way that brain works to judge the plausibility of the assumption that if the brain suddenly

got better, bigger, you'd get this. And there's pretty good reason to doubt it. One of the people in the early 50s, a couple of graduate students who were talking about these problems, Eric Lenneberg, who went on to sort of found modern biology of language. One of the things that he studied was what we call pathologies. Like one of the topics he studied was called nanocephalic dwarfs. That is, humans who have extremely small cortex. And he found cases of an extremely small cortex with perfect language capacity. Maybe there's more recent… You probably know [*Saying something to Angela Friederici*]. You should probably be asking Angela about this stuff.

Friederici: Well, I think with this final question and final answer, we come to an end, we have to come to an end. It's, I know, it's an unfinished business, there are much more questions also in the audience, I saw hands being raised. Well, you think we should have one final question from a female asker?

Chomsky: This is a very serious social problem… which is very, very hard to overcome…

Friederici: You were raising your hand already several times [*pointing and one woman*]

Woman: I was just wondering… language is innate to humans. Language developed only about 75 000 years ago which is left relatively late in the history of humans. When we think that the split off of our nearest relatives was about six million years ago, then isn't there a more fundamental innate feature that separates us from animals? Something like shared inventionality or…?

Chomsky: That could be, yeah. You know, in fact there are some – if you look – are you a biologist? [*asking the woman*] [No] – Well, if you take a look at

contemporary evolutionary theory, some very interesting things have been discovered. One of them is conservation of physical structures and capacities. In fact there is now a theory which is taken seriously that there is a universal genome, which holds for bacteria up to humans. It's basically all the same. Everything we now find developed suddenly in and around the Cambrian explosion and then changes from say bacteria to other species and humans and so on are just kind of like minor modifications of it. I don't know if it's true, but it's taken seriously There's enough information about deep conservation so that you have to take even extreme theories like that seriously. And it might turn out to be true that there's something deeply embedded in organisms that has the latent capacity under certain conditions of coming out this way. These are all really new topics of the last 20 or 30 years in biology. And it's pretty exciting work, and new things are happening all the time. So, if you keep up with contemporary work and evolutionary theory, every time you open a journal, you find more examples like this. So it's not inconceivable, but all we know is that the behavioral effect is very recent and there is nothing two hundred thousand years ago that even begins to look like what happened with this particular strain of hominids that we came from. The closest, not quite - almost existing -- species is Neanderthal. They were around until maybe thirty thousand years ago. And lived together with homo sapiens in the same places. And Neanderthal were all over the place. They spread all over the world, and they left plenty of artifacts. So, there's a lot of study of what you can attribute to Neanderthal's cognitive capacities. And it turns out to be quite interesting, so for example Neanderthal were fantastic tool makers. I mean they could make tools that humans today can't make without advanced technology. They really had extremely sophisticated tools, but it turns out that the tools were *identical* wherever they were. You know, so they could have been anywhere in the world if you look at the artifacts going back hundreds of thousands of years, the tools are the same. Extremely intricate but the same. If you take a look at humans

from the point of this so-called great leap forward – Jared Diamond's phrase -- it turns out all of a sudden tool-making becomes extremely complex and varied and there's a lot of creativity in the use of the things and so on. So something seems to have happened there. Maybe it's deep conservation. But you know, maybe in important ways we're all like bacteria.

Angela Friederici: Well Noam, thank you so much for putting the bits and pieces together for us from tool-making on, sort of going – and telling us what that could mean for the evolution of the human being and finally the evolution or the mutation of language. We can only thank you for coming, thank you for being with us, sharing your ideas with us and discussing with us. Thank you very much.

End