The Interesting Part Is What Is Not Conscious: An Interview with Noam Chomsky

Michael M. Schiffmann

The following is the transcription of an interview held on 23 January 2013 at Chomsky’s MIT office in Cambridge, MA, between Michael Schiffmann (MS, in italics) and Noam Chomsky (NC, in regular typeface).

OK, it’s January the 23rd, we are at the MIT in Cambridge, Massachusetts. This will be a 60 minute interview with Noam Chomsky on the sixty to sixty-five years of his work, and we will try to cover as many topics as possible. To start off with this, I should put this into a context. I first started to interview Noam Chomsky about this [i.e. the history of generative grammar] two and a half years ago, right here at the MIT, and inadvertently, this grew into a whole series, and today’s interview is meant to be the end of the series, but not, hopefully, the end of our talks. [Both laugh.] Well, as I see it, and that’s a central part of the research project on your work I’m working on, there are, among many others, several red threads that run through your work, and that would be, first, the quest for simplicity in scientific description, and as we will see, that has several aspects, then the question of abstractness, which we will see in comparison to what went on before and what you started to work with. A closely related question that came to the forefront later was locality, local relations in mental computations. Fourth, the question of biolinguistics, meaning that language can be, and is seen by you, as a biological object in the final analysis, and also, that would be the fifth point, everything you did has always been developed in close collaboration with other people.

So it’s not, we are not simply talking about the work of Noam Chomsky, but it’s a collaborative effort. Starting in 1946, I remember from my previous interviews that that is actually the period when you got to know who would become your teacher later on, Zellig Harris. And one of the first things you did was to read the galleys for his best-known work, Methods in Structural Linguistics (Harris 1951). There is another anecdote that I just saw in the morning, when for the first time I saw that Barcelona—I think it was in Spain somewhere in November—talk,1 when you said that another motive, apart from meeting Harris, for going into linguistics, was that you discovered that the Bible, the first words of the Bible had been mistranslated. Can you—maybe that’s a good point to start.

Thanks go to Philadelphia filmmaker Ted Passon, who acted as the sound- and cameraman. Additional thanks go to Chomsky’s personal assistant Bev Stohl, who helped to arrange the interview, and to Noam Chomsky himself, who took the trouble to review and correct my transcript of our conversation.

1 Actually, the video, published just one day before the interview, was of a talk Chomsky gave at Princeton University on the invitation of Robert Freidin. See http://www.youtube.com/watch?v=Rgd8BnZ2- iw&list=PL5affof_lboya5kn0H7YbeQITE5oX9AS.
[Chuckles] Actually, I read the galleys before I had started taking linguistic courses. I had just met Harris, we talked about the field, and he said, you want to learn about the field, why don’t you proofread my book.

_Nice method!_

So I proofread the galleys for him. I think it’s probably mentioned in the acknowledgements,² but that was my introduction to the field. This was different. I was studying Arabic; that was one of the few fields that I was interested in, and the professor was a very distinguished Arabist and also a wonderful human being. He was an Italian, Giorgio Levi de la Vida, he was an anti-Fascist émigré.³ We got to know each other pretty well later, but he pointed out to me something, just in conversation, something about Hebrew—I knew Hebrew reasonably well, and knew the Bible, he pointed out to me—I forget the context—that the first few words of the Bible were misvocalized, you know, they were—the original text of the Bible, or the texts, came down without [vowels], it had just consonants. Hebrew, you know, and Arabic, you know, are missing the vowels; they’re extra. The vowels were put in about the eighth century by the Masoretes,⁴ and they just made a mistake in putting in the vowels. And the phrase that appears is completely ungrammatical. And the translations are wrong.

_Is the meaning also—_

The meaning, too—I mean, it doesn’t mean anything, literally, but it’s kind of been reinterpreted so that it means something. And it’s not—if you get it correct, it doesn’t change the meaning enormously, but I was struck by the fact that for twelve centuries at that point, the first words of something that everybody knows were mistranslated and misvocalized, and nobody had noticed it! And that struck me as meaning: Well, maybe there is something interesting in this field that you want to figure out! [Laughs.]

_Right! Well, I guess the morale would be, there is—we think we know so much, but actually, so little is known.

Yeah, we don’t pay attention to what we don’t know.

_Exactly. OK, so you started with Harris, and we covered this a lot, the basic idea really was at the time, in structural linguistics, not only with Harris, to look at things that were supposed to be ‘there’ in the world and to classify them._

Yes, you study the corpus, you study the corpus of material, when you’re a student, let’s say in the undergraduate stage, in the late forties when I was a student, you were taught field methods, how to take information from an informant. So it would

---

² It indeed is. See Harris (1951: v).

³ Giorgio Levi Della Vida (1886–1967) was a Jewish linguist active in the anti-Fascist movement in his native Italy. After the introduction of anti-Jewish race laws in Italy in 1939, he emigrated to the United States, from where he returned to Italy only in the late 1940s. His main areas were Semitic languages, in particular Hebrew and Arabic. He also worked in the realm of Near-Eastern history and culture.

⁴ The Masoretes were Jewish scribes between the 5th and 10th century and, among other things, devised a notational system for putting in the vowels missing in the original text of the Bible.
be typically a Native American speaker of Choctaw or something, and you have methods for getting information and data from him about the language. And then you organize that data. And that’s a grammar, you know.

You go from phone to phoneme, and on to morpheme, and on to syntax—
You have a rigorous series of steps, you can’t mix levels, you can’t use syntactic information to analyze the sounds—
Yes, which you later on did in 1956, in that article—^5
Well, in 1946, because as soon as I started working myself, I saw this doesn’t work, you know.

OK. But you say you immediately saw there is something fishy about that, though that would begin the phase in your life, in your linguistic life, that you have later on called schizophrenic.の On the one hand, you tried these methods—
Well, I assumed—you know, Harris was an extremely impressive person, did a lot of interesting work and so on, I just assumed it’s gotta be right. Thus one part of my life for several years was trying to fix up the procedures and analyses that wouldn’t work, because when you looked at them closely, they didn’t really work and so on, so I tried to sharpen them up, and overcome difficulties, and I continued with that until about 1953. On the other hand, there was something else, which was kind of like a private hobby. I mean what happened is that we didn’t really study much linguistics as undergraduates, because Harris, Harris’s view was that the field was basically over. The Methods showed, and a lot of linguists believed this, once we have the methods of analysis, it’s all just routine from then on.

He had already begun the level, begun to move beyond the level of sentences into discourse—
The only things we studied—like, there aren’t, there weren’t—it was a funny department: I mean, there were a few students who were close to Harris; we met separately, and sometimes in his apartment in New York or something like that, and the things that were discussed were what became discourse analysis, he published a couple of articles on it in the early fifties, an effort to apply the methods of linguistics, assumed to be known, to more complex discourses. Essays, articles, chemical papers for the purpose of abstracting, and so on and so

---

^5 See Chomsky et al. (1956).

^6 In conversations with the French linguist Mitsou Ronat in January 1976, Chomsky described his situation between 1947 and 1953 in the following way:
I did not think myself that I was doing linguistics. In a sense I was completely schizophrenic at that time. I still thought that the approach of American structural linguistics was essentially right. As I told you, I spend a great deal of time trying to improve and to formalize discovery procedures, in order to overcome their obvious defects. But once they were made precise, they led manifestly to the wrong results. Still, for quite a long time I thought that the mistake was mine, due to wrong formulations. In 1953 I published an article in the Journal of Symbolic Logic in which I tried to develop a discovery procedure that I hoped might be the basis for something that would really work. That, for me, was real linguistics. Chomsky (1979: 130–131)

^7 See Harris (1952a, 1952b).
forth. That’s basically all we studied. A lot of it was political. Stuff I could tell you about.

Yeah, I know.

Because that was the real interest that brought a lot of us together, but so, efforts to detect ideological factors by looking at texts from a structural point of view, things like that. There was a—I had to do an undergraduate thesis, an honor’s thesis, you know—

And that undergraduate thesis had nothing to do with that, it was different.

No, it was linguistics. And Harris suggested as a topic that I, since I knew it, that I do a structural analysis of Modern Hebrew, which I kind of—I wasn’t fluent, but I knew it pretty well, and so, OK, so I went ahead and did what I was supposed to do: I found an informant, a native speaker and went through the procedures, and then I got the phonetics, and on and on. After a couple of weeks of this, I realized that this is totally ridiculous. The only thing that I’m learning is the phonetics, which I don’t care about. [For] everything else, I know the answers he is going to give me. Why am I bothering with the questions?

I see.

So I dropped the informant altogether, and just started doing what looked like it made sense, which was in fact a generative grammar. So let’s just, let’s construct a generative system, which will determine the nature of the expressions of the language. I don’t need the informant for this, I can—

Can we briefly move into the syntactic component of this, because there were like twelve pages of syntax in that first work of yours. And it was highly formal, as you several times mentioned, well, probably nobody ever looked at it because nobody understood it—

One person looked at it.

Yes, Henry Hoenigswald.

We talked about this.

Yes, and part of that formalization was that it had recursion in it.

This it had to automatically, because it’s sort of like a truism that language is unbounded. And it’s kind of interesting that later, when I studied the history of the field, I discovered that this had barely ever been mentioned. I can run through cases—

I mean, of course it’s kind of obvious that it’s recursive, but nobody ever mentioned it—

Nobody mentioned it was infinite! I mean, few people. Scattered cases where it

---

8. This refers to the final, published version of what started as Chomsky’s 1949 BA thesis. An extended and revised version of that work served as his 1951 MA thesis. This thesis was once more revised in the fall of 1951 after discussions between Chomsky and his friend Ye-hoshua Bar-Hillel, revisions that were completed in December 1951. This is the version that was finally published almost three decades later as Morphophonemics of Modern Hebrew (Chomsky 1951/1979), from now on, MMH.

was. They’re interesting.

Yes. And here it was right in the first rule: A sentence is something which can be re writ-
ten as something that also contains a sentence, and it goes on from there. That would be the direct recursive rule, where you have the S on the left side of the arrow and —

Well, the system is recursive; there is no particular rule that’s necessarily recursive. The set of rules gives a recursive function. It’s analogous to — I should say at the same time I was studying recursive function theory and mathematics and so on, and it sort of all fit together. That provided the formal understanding of what a recursive system is.

Exactly, I mean the progress that was made at the time, that was that things were laid out in the open, that had kind of been lingering in the background of people’s mind, but never [been] formulated clearly.

Well, you know, Euclidian geometry was kind of a computational system. It didn’t work exactly, in fact it wasn’t even formalized until the turn of the 19th century. But then the whole concept of algorithms, computations, and so on, that had really been, by the 1940s, that was pretty well understood with, very well understood with Turing, and Church, and Kleene, and other great logicians.

So in Morphophonemics, you had this expressly recursive system, with the S on one side of the arrow, and also on the other side, and then you had NPs that could contain PPs, and the PPs in turn could contain NPs —

Yeah.

— that would be another example for recursion, but compared to what you later did, this was not transformational syntax.

No, this was pre-transformational, yeah.

So the transformational work started around here, in Cambridge, right?

In the early 50s.

---

10 The first syntactic rule in MMH “S1. Sentence → Elementary sentence <Connective + Sen-
tence>”, accompanied by the comment: “S1 must be reapplied until ‘sentence’ is eliminated” (p. 12). In the conversation between NC and MS, ‘S’ is used for ‘Sentence’.

11 As it stands, this is not really correct, at least for the presentation in MMH. Since the syntactic component of Hebrew in this work is only a sketch, and outlined only insofar as it is necessary as input for the morphology, it contains no ‘indirect’ recursion (i.e. recursion resulting from more than a single rule), as could, for example, be exemplified with the set \{PP → P + NP, NP → N + PP, NP → N\} of rules whose repeated application could generate strings such as \{for [NP prisoners] in [NP prisons] within [NP prisons]\}]. Omitting irrelevant detail, rule S5 in MMH can be stripped down to PP → P + NP, but the following rule (actually a collection of rules collapsed into one) S6 expanding NPs does not contain the second step necessary for indirect recursion, namely NP → N + PP or some equivalent. As far as I can see, there are also no other syntactic rules in MMH that would yield indirect recursion. This system thus only provides for direct recursion, and what is more, the only rule doing so is the one mentioned in the previous note, where the symbol ‘Sentence’ appears on both sides of the arrow. Also underlining the sketchy character of the presentation is the fact that the only ‘Connectives’ spelt out in rule S2 are coordinating conjunctions, so that even sentence recursion is limited to coordination. The transformations introduced in the LSLT/SS system (Chomsky 1955, 1957) kicked recursion out of the phrase structure component of the grammar entirely, at least until it was reintroduced in Aspects (Chomsky 1965).
When you came to Harvard, as a fellow.

Yeah. Incidentally, the syntax was pretty rudimentary. That work you’re talking about was about 95% about the mapping from the syntactic structures to the sounds.

That’s pretty clear, and was in fact laid out in the introduction to the booklet: “I’m doing this just to give a base to do the other things.” So, one of the things that I remember you telling me that you did was around 1952 when you worked with that famous auxiliary transformation, with the English auxiliary system. Can you give a brief description?

Well, I was struck by the fact—in the descriptive grammars of English, they ran through the paradigms of the auxiliary system, you know: ‘eat’, ‘was eaten’, ‘has eaten’, ‘will have eaten’, and so on and so forth. The forms were just listed, and that was it. But if you looked at the forms, there were clearly some patterns behind them, like when you have ‘have’, it was normally followed by something which ended with an ‘-en’ element or some other morphological component; ‘be’ was going to be followed by ‘-ing’, and so on, and they appeared in certain ways. And there was a very simple way to describe all of this, namely, just to say that, the way it was described at that point was: The real element is not ‘be’ and then a separate form ‘reading’, but ‘be-ing’, and ‘read’, and then the ‘-ing’ moves over and attaches to the ‘read’—or the other way around! But that’s a transformation—

And in that case one element determines the form of the next, in that the affix hops from that element on to the next one…

Yeah, what was later on called ‘affix hop’, years later, by Haj Ross.12 But—

He had a gift for those formulations.

He liked those terms, yeah. A lot of his terms were—I didn’t actually like them, I thought they were kind of frivolous, frankly, [laughs] but I use them—

Like ‘island’ and so on…

—but now it’s called ‘affix hop’.

OK.

But—not my term. But the point is, it wasn’t a phrase structure rule. And then, when you begin to think about it, if you take, say, “the sandwich was eaten”, not only you have this affix hop, what you also have the subject not in the position where it gets its semantic role. Well, and if you think about that, that should be another transformation. Now, Harris’s descriptions had—at that time, he was working on co-occurrence, and he tried to, he gave a descriptive argument to try

---

12 Haj Ross, born in 1938, studied at and got his Ph.D. from MIT (Ross 1967). Despite its great significance for linguistic theorizing, it was only published 20 years later (Ross 1986). Ross has worked, and continues to work, on a huge variety of linguistic phenomena. He was part of the short-lived linguistic current called ‘Generative Semantics’, and, using a metaphorically very gifted language, he has coined a number of technical terms such as ‘scrambling’ (the shuffling around of syntactic constituents from their original position to other positions not canonically designed as end points of such shuffling), ‘islands’ (syntactic structures that constituents are more or less unable to move out of), ‘pied-piping’ (constructions in which a question word such as whose ‘carries along’ the phrase which it is a part of), and more.
to show that passives had co-occurrence relations with actives, but it didn’t really work quite properly. And I thought then, as I say, for a while, for a long time I tried to work it out, but then I realized that this is not the right way to do it. The right way is that there is just an abstract underlying structure, and there are various kinds of rules applied to it, one of them being the passive.

And that’s the really important thing, you have something abstract from which what you see, finally, or what you hear, is derived, and that’s basically the difference between transformations in your work and Harris’s.

Yes, and that means that you also abandon the separation of levels. And there was a lot of artificiality that came along with the separation of levels, I mean simple things, like take the words ‘writer’ and ‘rider’ in English.\footnote{See, e.g., Chomsky (2010a: 4), Halle (2010: 18, fn. 1), and sources quoted there.}

The famous example, yes.

Famous example. For structural linguists, who did work on this, phonetically speaking, say, in Standard American English, the phonetic difference between ‘writer’ and ‘rider’ is in the vowel.\footnote{In (one) phonetic notation, the pronunciation is [rʌɪdər] and [ra:ɪdər], respectively. The difference lies in the first vowel, not, as intuition and the spelling of the items might lead one to assume, in the medial consonant.} The medial consonant is identical for the two. And since you’re not allowed to mix levels, the description—that’s the way the description was—intuitively that’s obviously crazy. There is a word ‘write’, and there is a word ‘ride’, and the vowel change depends on the consonant, and then the consonant is neutralized.\footnote{See previous note.}

And once more the truth of the matter comes from things that you cannot see, because what you see and hear is derived from some computation that’s—

Well, the point is you weren’t allowed to give this description, because it violated the separation of levels. You had to know the morpheme—

Yes.

Harris had various gimmicks just to try to get around this, but they didn’t work, and my feeling was: Let’s do it a way that works, that’s the way I had done it with Hebrew with just ordered rules.

And maybe the most extreme example maybe of mixing of levels would be, on the one hand, you have syntactic structure, and then you have things like contour and stress assignment—

Yeah.

So you have the highest structure and the lowest structure, and then [gestures up and down to connect the two].

That’s something that Morris Halle and I worked on in, well, around 1953, ’54, and we published an article in ’55. One of the major topics at that time in American structural linguistics was pitch and stress.
Yes, *Trager and Smith*.\(^{16}\)

And there was a system, which was supposed to be an overall system for English, which had four stresses, four pitches, and then you just describe things within that system. But Morris made an interesting observation to me when we were working together. He pointed out that Roman Jakobson makes every imaginable mistake when, in the phonetics of English, but he never gets the pitch and stress contours wrong.

Yes, that’s the one thing he gets right.

So we figured: Look, there’s gotta be some rule system behind this, and when we took a look at it, it turned out that you can predict pretty complex stress and pitch contours just by regarding them as the reflexes of the syntactic structure with cyclic rules that generate bigger things from smaller things. Of course all of that is totally outside the framework of procedural analysis, because you’re using abstract syntactic structure to derive phonetic facts, so it was going to be kind of like incomprehensible in that structuralist framework. But it looked right, and we, Morris and I, then pursued this for years, and finally, in *Sound Pattern of English*, so it’s a long discussion, and it goes on from there. But that was basically the idea, that it’s in the morphophonemics—incidentally, I later learned, only much later, that there is a historical tradition, which goes back to classical India, the Paninian tradition, I may have mentioned this before—\(^{17}\)

Yes, you did.

Yeah, OK.

*I have a very little question about the auxiliary system as you described it, and it’s really a super-technical question, but, when I with a group of students read Syntactic Structures,\(^{18}\) we had this rule for the auxiliary system: C, M, then ‘have + en’, then ‘be + ing’—what does the ‘C’ stand for? Where does that ‘C’ come from? Is it ‘Concord’ or something like that, do you remember that?*

[Laughs] Oh, that’s sixty years ago…

[Laughs] I asked people at home to give me questions along, and I thought this was a good opportunity to nail this one down.

Absolutely. I’m really sorry, I would have to recover pretty old memories to find out what was ‘C’. It’s not the ‘C’ that later became standard, the ‘Complementizer’.

*Of course, absolutely not, that’s pretty clear.*

It had to do somehow with tense, but I forgot why I picked ‘C’…

---

\(^{16}\) Henry Lee Smith and George L. Trager were important structuralist linguists and authors (jointly and individually) of works on the structure of English. In the 1950s, their theory about intonation, and stress assignment was dominant in the US. Chomsky discussed this particular issue as well as phonology in general and even more general issues with Smith at the *Third Texas Conference on Problems of Linguistic Analysis*, May 9–12, 1958 (Hill 1962). For some additional remarks on this, see Chomsky (2010b: 2).

\(^{17}\) See Chomsky (2010b: 19-20) and Chomsky 2010a).

\(^{18}\) See Chomsky (1957).
The only thing that sprang to my mind was ‘Concord’—

Well, maybe it was ‘Concord’, but it was basically the tense morpheme.19

Yes, I see. Let me get to the biolinguistic question at the time. The Harrisian tradition was not realist; the idea was not what we are describing here as truth about the world or the language or whatever—when did you start, when did you start and when did Morris start and when did your friend Eric Lenneberg start to develop the different conception? Would that be around that time, about 1951—

Well, the three of us met in 1951. We were all—Morris was here, but Eric and I, we were grad students together, and we met pretty quickly and kind of hit it off, we were all friends and we started thinking, talking about these topics. None of us—one thing we had in common was, none of us believed in structural linguistics, or anything that was going on. But—you have to remember this was a period when strict behaviorism was absolute dogma, in philosophy, in psychology, and we didn’t believe in that either.

So right from the start.

Right from the start. So we started reading, you know, together, a lot of the European ethological literature. Morris and Eric could read it in the original, it’s German, and I read it in translation, but it was just excluded from American psychology. Almost, I mean there was—you could find pieces in Comparative Zoology,20 in Comparative Psychology,21 but not mainstream psychology. I mean, there are some obvious points about language, I mean, it’s just like about as close to a truism as you can get: A language, say, my language, is a property of me [points to himself], an individual! It’s not some abstract thing out in the universe somewhere.

Well, there are still people who claim that.

I mean, people claim that, it’s so crazy I won’t talk about it, I mean, this is kind of like as obvious as you can get. Your language is a property of you, you’re a biological creature, so it’s a property of your biology.

If somebody knocks me on the head hard enough, it’s gone.

Yes, and in fact, well, furthermore, it’s a property mostly of your brain, because when your foot’s cut off, you don’t lose your language, if your head is cut off, you do lose your language. But that’s something that’s too obvious to discuss. So it’s a biological property of an individual—now, you can study language in a different way if you chose.

Right.

---

19 In his final editing, NC adds: “Actually, I suspect it might have been ‘component’, borrowed from the Harrisian concept long component—an inflectional element that spreads over a sequence of elements” (May 2014).

20 The Zoologischer Anzeiger: A Journal of Comparative Zoology, one of the oldest German zoological journals, is a scientific quarterly that has appeared since 1878; it is now published by Elsevier.

21 The Journal of Comparative Psychology has been published by the American Psychological Association since 1921.
But at least one way of studying it, and in fact a central way, is as a biological property of an individual. Well, and let’s think of it as a biological property, what do biological systems have, like the visual system? They have a genetic basis, they grow in particular ways, they’re used in particular ways, they evolve, but that was way too far off to even think about... And that’s biolinguistics!

And that would mean that something that you mentioned often, a quote from Martin Joos, that languages vary without any limits,\(^{22}\) can’t be right.

It can’t be right, because there would be no way to acquire any of them. And I quote him, because he said it clearly, but it was everyone’s view. Or essentially without limits, let us assume there are some constraints—

What occurred to me recently was that behaviorism in other realms said exactly the same thing! For example when, I believe it was Watson, one of the co-founders, who said, “Give me some child, and I have my behaviorist methods, and I’ll make a criminal or a saint, I can do whatever”—

Yes, but it was the same actually in theoretical biology. It was, not really by everyone, but it was generally assumed, I’ve quoted some of this, too, that organisms can vary—

—without limits—

Almost without limits. You know, when people say, virtually infinitely, which means, just about anything. It’s interesting that in biology over the years it’s been recognized that in fact, the limits are very narrow.

Yeah. As you recently wrote or quoted, it’s just variations of one organism since the Cambrian explosion.\(^{23}\)

Well that’s actually been suggested seriously, as a proposal—

Right. There are other things that you can clearly see, and you mentioned that in another article, for example, the kind of surprising fact that animals rarely use metals in any significant sense, and animals don’t have wheels, even though it might be helpful to have ones!\(^{24}\)

[Both chuckle.] Yeah. Somebody pointed it out to me once, a biologist, but there’s apparently a certain kind of bacteria that have something like wheels —

OK! [Again, laughter.]

Not to get around, just—it’s things that spin, you know...

*But the basic point really is, it can’t be limitless.*

---

\(^{22}\) See Joos (1966: 96), citing the ‘Boas Tradition’, quoted (from an earlier edition) in Chomsky (1968/2005: 68, n. 12) and many of Chomsky’s speeches and writings subsequent to the first edition of this work.

\(^{23}\) See Berwick & Chomsky (2011), where the biochemist Michael Sherman is quoted with the suggestion that a “Universal Genome that encodes all major developmental programs essential for various phyla of Metazoa emerged in a unicellular or a primitive multicellular organism shortly before the Cambrian period” (p. 23).

\(^{24}\) Actually, two different sources, namely Chomsky (2002: 103) for the use of metals and Berwick & Chomsky (2011: 22) for the impossibility of using wheels.
You’ll do with, you’ll use the tools that are available to you. Actually, the most striking—things like this affix hop that you mentioned, it was only really one example. I mean, as soon as you started looking at, trying to actually generate the properties of the language, the actual properties, you quickly found that there are some very puzzling phenomena, things were one way, and not some other way, in which they could work. And this was never even noticed, you know. One of the simplest cases, which is still being debated, hotly debated—

_I know, yes_—

—is structure dependence!

_Sure, I know what you are getting at._

Well, that’s the kind—

_Structure dependence by the way also gives in my view, structure dependence in my view also gives a very interesting example for locality. Because if you have something like the sentences you quoted in your work in the fifties and also together with George Miller, with multiple dependencies, you have an ‘if’ here in the sentence, and in principle, the ‘then’ that comes after it might be a million words after it._

And once you move from linear order to hierarchical structure, it’s immediately adjacent, so the relationship becomes local, pure and simple, like in these examples, “Is the man who is in the garden happy?”.

It’s a little bit like affix hop, really if they’re in a single unit, and then they separate.

_Right._

But the structure dependence is much deeper than that. The examples that have usually been used in the literature involve movement, like auxiliary inversion, and that has misled a lot of people. There are all kinds of pointless efforts to find other ways of describing auxiliary inversion, but the very simple examples that don’t have any of these properties, like just take adverb interpretation, so suppose you have some simple sentence, one that I’ve used is “Eagles that fly swim.”

And suppose you put ‘_instinctively_’ in front of it. Which verb does ‘_instinctively_’ go with? “Eagles that fly swim.”

_Yeah, that’s absolutely clear from the start._

It’s perfectly obvious that it’s ‘_swim_’, but _why_? I mean, the relation to ‘_fly_’, first of
all, is more natural; they do fly instinctively, but also, it’s local, it’s a local—
In the interpretation you even go against the pragmatics of the expression.
Against the pragmatics, and also against ease of computation!
Yes.
Because the local relation, the one that’s easy to compute, is with the first verb!
But you don’t, you find—it’s also a locality relation, but it’s structural locality. It’s
the structurally closest one, not the linearly closest one. And that’s much harder
to compute!
And you go against this, because language happens to be set up this way—
Well, the real answer to this, I think, which is just becoming clearer now, is that
order doesn’t exist in the internal computation. So, like when a child—it’s just not
available, which has a lot of consequences, we could talk about it. But if order is
part of the externalization, it’s kind of a reflex of the sensorimotor system. It’s
outside, which implies that other semantic and syntactic properties also
shouldn’t depend on order. And there’s pretty good evidence [that] that’s true;
they depend on hierarchy, not on order. Order seems to be extrinsic to language.
And from that a lot of other things follow.
Right, we talked about that in Cologne—
We did.
—for example, that islands and a lot of other, similar stuff, that you can, that in principle
certain thoughts are fine, but you cannot express them in the externalized language—
that that follows from that.
Well another thing—yes. All of this converges on—has a lot of syntactic and se-
matic consequences, but also much more general ones. Because if order is ex-
trinsic to language, it then follows that every use of external language is also ex-
trinsic, in particular the use of language for communication, which has got to be
secondary, contrary to dogma, another dogma. Anyhow a lot of these things we
were kind of talking about—I’m going back to the fifties—we were kind of play-
ing with, but Eric, he finished his linguistics and went off to medical school and
ended up pretty much founding the field of biology of language, and we con-
tinued to work with each other for some years.
I was wondering about a point we talked about in Cologne, because externalization pre-
vents you from expressing certain thoughts, and as long as you talk to yourself internal-
ly, you are still using the externalized language, and that would mean that in that con-
text it would be the same. What about unconscious thoughts? Would that remove the
barriers to these thoughts?
Well, when we do what we call ‘talking to ourselves’—this is not investigated,
but it could be investigated and it should be, but if you just introspect about it

28 To be precise, in the Cologne interview the issue was not order as such, but the island phe-
nomena mentioned by MS in the next few sentences. See Chomsky (2011).
29 Ibid.
carefully, when you think you’re talking to yourself, what’s actually happening is [that] fragments are passing through your consciousness, not sentences.

Yes.

What’s hitting consciousness is bits and pieces, and then you can formulate an expression, using them, and that’s talking to yourself, but the real thinking is beyond the level of consciousness. It’s something that’s going on in there using this whole system; occasionally, bits and pieces of it hit the level of consciousness, and that’s what people study, what’s conscious, but the interesting part is what’s not conscious. In fact, there’s other information—there are other studies now showing something similar. For example, it’s been demonstrated that if you carry out an action, you know, willed action, a tiny time before the action is—you think you’re deciding to do it, there is already neural activity in the motor areas, which means that the actual decision is unconscious.

Yes, we talked about it in Cologne, and the absurd fact that people used that as an argument against free will.30

Free will, well, no, it’s not—

It makes no sense.

It makes no sense, and it’s kind of interesting [that] in the history of thought, the concept of unconsciousness has barely entered. I mean, even in Freud, he talks about the unconscious, but you can bring it to consciousness—

Yes, that was the basic idea.

That’s the whole point of psychotherapy.

Yes, of course, I know, and it’s a big philosophical point which we have no time for… Let’s hop back into the 1950s.

OK.

I guess when you wrote that big work of yours, Logical Structure of Linguistic Theory,31 what was clearly in the background was that languages can’t vary in limitless ways, otherwise, you wouldn’t have bothered to write about this logical structure and set up this system —

Well, the obvious reason why languages can’t vary in limitless ways is that there would be no way to learn any of them. I mean, this had to do with—it has to do with the impossibility of induction without a framework, and incidentally, that was one of the topics that Nelson Goodman—

I mean, that sentence alone is sufficient; you have to have a framework, otherwise you can induce nothing…

Which is—it was quite interesting at the time, because the other person32 I was
studying with was Nelson Goodman, who really argued that you didn’t need a framework! He tried to show in his new literature on induction, in *Fact, Fiction, and Forecast*, he posed fundamental problems for induction, and then argued that there were ways out of them by looking at the way things were used, but it didn’t work at all, you know.

yeah. And looking at the system as you set it up in LSLT and comparing it to Morphophonemics, you have the transformations, so it’s a two-part system; on the one hand, you have the phrase structure component out of which the kernel sentences grow, and then you have a transformational component, which serves two purposes, namely doing transformations within the kernel, and also, sticking the structures that underlie the kernel sentences, sticking them together. So that would be a different set-up from what you had in Morphophonemics.

Well, yes, Morphophonemics had only some, a few phrase structure rules, but in LSLT, it was assumed that, yes, there is a phrase structure component, but it generates abstract structures.

Of course!

And they have to be mapped onto what turns out to be a sequence of morphemes.

The phrase structure component doesn’t create the kernel sentences, but the structures underlying the kernel sentences…

It generates only the structures underlying them. And at that point, I also assumed that there was a finite number of them. I assumed that the recursive character of the language was in the transformational component, in fact in generalized transformations.

That is the point I was wanting to make. Yeah, and I see a very interesting parallel here in having sentences that are stuck together by generalized transformations, in terms of things that now crop up again and again through ‘barriers’, and now ‘phases’. The whole thing seems very similar to me. Because you have some localized things, within that very, very restricted phrase-structural component, and of course you have very local

which appeared in the same journal that was also the venue for Chomsky’s first published article (Chomsky 1953).

33 See Goodman (1955).
34 The concept of ‘kernel sentences’ goes back to Harris; they are, roughly, simple declarative active sentences.
35 This concept was developed in Chomsky’s (1986) monograph carrying the same title. At their core, the concept of syntactic barriers deals with the question of how far, and in what steps, syntactic elements can move. With it, Chomsky tried to unify two independent constraints on movement, namely Subjacency and the Empty Category Principle (ECP). For some informal remarks that might elucidate both the notions of subjacency and ECP and the unifying idea of barriers, see Rizzi (2012) and footnotes there.
36 ‘Phases’ are syntactic units roughly corresponding to verb phrases accompanied by the verb’s assignment of semantic role to its subject and object, full finite sentences, and probably determiner phrases in one form or another. The idea here is again that the rules of human grammar are first applied in the smallest phase, which can then be ‘forgotten’, then the next larger one etc. A simple and non-technical presentation of the idea can be found in Larson (2010: chaps. 26–27).
operations now in phase theory. So—

Well, it goes beyond that. Take a look at the work I and others have been doing for roughly twenty years, called the Minimalist Program. It bars, it goes back in a lot of ways to something like generalized transformations. For one thing, there’s no phrase structure grammar. Phrase structure grammar was basically eliminated in the 1960s. Phrase structure grammar involved a tremendous number of stipulations, like, why do we say “V → V”, [I mean,] “VP → V + NP”, why don’t we say “NP → V + NP”, you know? It’s just stipulated, so endocentricity is stipulated, where the phrase is compared against the head. There is massive stipulation we want to get rid of. It was more or less gotten rid of by X-bar-theory, and that has its own stipulations—

I think at the time it was not formulated that way, but for example, if one looks at the work that comes after Aspects of the Theory of Syntax, and one scans the work, your work for example, for phrase structure rules, you will find them as an expository device, for example, the two-part phrase structure of X-bar-theory; it’s nothing more than that, and the meaning is completely different, in that things are flipped around—phrase structure is top down, and X-bar is also [makes a gesture from below]

It’s projected from the lexical items, the smallest atoms. Actually, there is new work I have on this, which suggests a slightly different way of looking at it, but I think you can get around, eliminate a lot of stipulative character of X-bar theory as well, but when you finish this, you will really do a lot of parallel computation, so if you construct not even—I mean any sentence, like “The man saw the boy”, let’s say. ‘The man’ and ‘the boy’ are being constructed in parallel, ‘see the boy’ is being constructed out of ‘see’ and ‘the boy’, and then ‘the man’ is being added to that, that’s all parallel computation, which kind of has the spirit of generalized transformations in a way, so if the only combinatorial operation is the absolutely simplest one, namely, just Merge two things to make a new thing, then you’re going to have parallel computation, extensively, richer even than generalized transformations.

Yes, I see. Let me, since our time is already compressing very rapidly, let me get back to the 1960s about things that we haven’t talked about yet so much. One thing that really

---

37 Reference here is to the introduction of X-bar theory in Chomsky (1972). From that time on, X-bar theory in its various formulations, started to replace the previous multitude of phrase structure rules that had characterized generative grammar from its beginnings and even more so after the recursive component, the part of the grammar that enabled infinitely long sentences, had been assigned to phrase structure rules instead of generalized transformations in Chomsky’s (1965) classical Aspects of the Theory of Syntax in what henceforth was often called the ‘Standard Theory’.

38 In his final 2014 corrections, NC says: “Might be better to delete this [namely, ‘so endocentricity is stipulated, where the phrase is compared against the head’]. Endocentricity really comes in with X-bar theory, not phrase structure grammar.” I have kept the passage in order to illustrate how the attempt to get rid of one set of stipulations can inadvertently lead to new stipulations that may come under scrutiny only much later.

39 Simplified to the core structures, the two residual phrase structure ‘rules’ in Chomsky’s (1986; 3) Barriers are X’ = X + XP and XP = XP + X’. Presumably, it is hardly an accident that Chomsky doesn’t use the arrow ‘→’, meaning ‘consists of’, to present the relation, but rather the more neutral symbol ‘=’, which suggests a two-way relation at the very least.
became foregrounded in the 1960s was what you just mentioned, namely the question of language acquisition. It was always there, in that you had to have an inductive system, but starting from the 1960s it was also applied in research, including the research of your wife, including the research of Lila Gleitman and so on and so forth. What do you think are some of the important things that happened in terms of this, of language acquisition research in the 1960s?

Well, actually the first—before that, language acquisition studies of course had taken place, but they were very descriptive. Experimental work began, and in order to study how a system is acquired, any system, you have to know what the system is! Now the earlier work didn’t really ask: “What’s the system that’s being acquired?” They just looked at, you know you have to, children learn simple words, and then they learn two words, and then they learn three words, and so on. But as soon as you ask, “What system is it that’s being acquired?”, you ask different questions. So for example take Lila’s work, Lila Gleitman. She was asking questions like, what does a child understand when it’s producing only two words. Well, it turns out it understands much more complex structures. With the right kind of experimentation, you can show that—

_In that talk in Spain in November last year, you said, well, research tends to show that pretty much everything about the language is in place maybe at the age of two._

That’s—there is work that suggests that. Actually, it’s my wife’s work. But for example, if you take cases like, there is the famous Helen Keller type case. Helen Keller was, she lost speech and hearing after around twenty months, and then she developed—

_An extreme case of the poverty of the stimulus!_

Extreme case. It looks like that. However, I think that’s a little misleading, because the evidence at least suggests that by twenty months she already knew the language! And everything else was extracted—

_I see the point. So which kind of shift the PoS argument because that, doing this in that period as a small baby is also extremely impressive, isn’t it?_

That just makes it even more dramatic. I mean the amount that’s known—if anything like this is correct, and my guess is that some of it will be shown to be correct, it’s hard to do—, that would mean that basically, it is pretty much like the growth of vision. I mean, the structure of the system is already fixed, and you just tinker some little bits and pieces of it.

_Another thing that it took people a long time to even think about, because it came so natural, of course, you grow up, you see, it’s normal, it’s there, but there is a lot to explain_

---

41 See e.g. C. Chomsky (1969). This was Carol Chomsky’s doctoral dissertation written under the supervision of Roman Jakobson.
42 For an incomplete account of Lila Gleitman’s many contributions, see her website at [http://www.psych.upenn.edu/~gleitman](http://www.psych.upenn.edu/~gleitman), her CV posted there, and also Gleitman (2013).
43 Should read Princeton, actually—see footnote 1 above.
44 Part of Carol Chomsky’s far too little known work beyond her seminal contribution (C. Chomsky 1969) is listed in the bibliography of a memorial volume (Piattelli-Palmarini & Berwick 2013).
The other thing that became prominent in your work in the 1960s, but I guess you’ve been very interested in that before, is the philosophy of language and the history of linguistics also. Can you tell me when you got into this?

Well, I mean, in the 40s and early 50s, I was mostly studying philosophy! With Quine, Goodman, and others, I didn’t really agree with it, but I started reading earlier philosophy, mostly on my own. And of course I met, I was in contact with people like John Austin, close contact, I knew the British analytic, contemporary analytic philosophers of the time, and by the late 50s, I was reading and thinking about earlier philosophical traditions.

Did you also occupy yourself — well, that’s a rhetorical question because I know you did — with Wittgenstein?

I read Wittgenstein, I didn’t meet him, everybody read Wittgenstein, and so, yes, I read Wittgenstein, and well, some things were there which I thought were interesting. So I had — but the actual course work I was taking, such as it was, was mostly philosophy. And though I didn’t really agree with Goodman, I was very close to him, and studied what he was doing. He was very, kind of like Harris, very impressive person, and interesting ideas, I thought they were wrong but very interesting, and stimulating. And then I got, I try to figure out how, but I think it was probably through, maybe Thomas Huxley, that I started finding out something about earlier ideas about language. They were in the philosophical tradition, I started reading Descartes, and Locke, and others, and —

Yeah, and you wrote that famous book Cartesian Linguistics. What made you pick that title?

Well, actually the title, I probably should have picked a different title, because most people who read it didn’t get beyond the title. So if you get to the first page, it points out that there is no such thing as Cartesian linguistics, there is just linguistics which turns out had borrowed and adapted certain, and developed certain ideas, which were also developed in Descartes, and used [in] and crucial for his philosophy, but that were developed in different ways —

What were these crucial points?

Well, the crucial point in Descartes, which was actually very significant in his own philosophy — he was basically a scientist, we call it philosophy —

Yes, that distinction didn’t exist.

His thought, his core research was to try to show that you could give a mechanistic description of everything in the world. That’s the ‘mechanical philosophy’ as it was called, what was assumed by Galileo, by Descartes, by Leibniz and Huygens, and Newton, you know, and this was, this is the core of early modern science. Try to show that the world is basically a machine. Of the kind that could be

---

designed by a skilled artisan. In fact, it was designed by a super-skilled artisan, you know, but meaning levers, and things pushing and pulling each other, and so on and so forth.

So the spontaneous kind of physics that you have, as an ordinary human being.

It’s so intuitive, isn’t it, and it was stimulated at the time by the fact that skilled artisans were doing things like that. Like if you walked in the gardens of Versailles, there were all kinds of complicated machines—they were stimulating in much the way computers stimulate the imagination of today. So the basic idea was, look, that’s what the world is, and we will show it. Galileo for example was frustrated to the end of his life because as he said he just couldn’t figure out how to give a mechanical explanation of the tides, or the motion of planets and so on, and Descartes thought he had a system which would do it, and that was his basic contribution. Nobody studies it these days, because it is known to be wrong, but then he did notice that there are things that don’t seem to have a mechanical explanation. And crucial for him was language. He pointed out in the Discourse on Method [that] every normal human being can carry out what I have sometimes called the 'creative use of language', you speak indefinitely, you can use indefinitely many sentences, new ones all the time, never heard before, other people understand them, they’re appropriate, you use them as appropriate to situations, but it’s not caused by situations, which is a crucial difference, and it’s not caused by internal mechanisms of your particular physical state: You could produce many different things, and it’s coherent and intelligible, and so on. These properties—

Descartes did not talk about rule systems or anything like that.

But that happened very quickly. That’s when you get what I was calling ‘Cartesian linguistics’. For Descartes, this was fundamental.

This is exactly what I wanted to know.

For Descartes, this is the core of his mind/body distinction. He says, well, we have body, you know, the physical thing, which is going to be a machine, but then there are these properties that you cannot account for mechanically, so we need—he’s a scientist—we need a new principle. And in the substance metaphysics of his day, that meant a new substance, so that’s res cogitans, and then the main scientific problems are, find out its properties and show how it’s linked to the other substance. You know, that’s where you get the speculations about the pineal gland and so on. But this is fundamental to Cartesian philosophy. Now if you look at the work of the period where the people were maybe influenced by

---

46 Namely, God.
47 See Descartes (1637/1998), one of the many accessible editions.
48 That is, the creative and non-deterministic use of language, a use that for Descartes was at the same time neither arbitrary nor random.
49 Descartes’s postulation of two substances, one thinking and free—res cogitans—, the other material and determined—the body—, posed the problem of how the two interacted with each other and what, if any, the particular locus of this interaction was. Descartes suspected that the pineal gland, located as it is right in the midst of that mysterious organ, the brain, was that locus and served as the means of that interaction.
Descartes, maybe not, some of the people were directly influenced by Descartes, like his kind of like junior associates, people like Cordemoy,50 he also pointed out that whatever this is, it is unique to humans, animals can’t do it, machines cannot do it…

He made that very explicit.

Very explicit. So you have—

And even human idiots have it and animals don’t.

Even idiots have it, but that led to immediate experimental questions, can we design experiments to determine if another entity who looks like us has this property or not.

Yeah, that would be the test. So that creature has a mind like us.

That’s where the basic work was done by people like Cordemoy and others. And that’s kind of like a litmus test for acidity: Do you have this property? It’s kind of similar to what people now call the Turing test, but much more significant, because this is a real scientific issue: We want to find an absolute—well, is there a test for this property? Then, when you get to the logicians and the linguists, especially at Port Royal, you know, the Port Royal grammarians where you have—they produced two great works, the Grammar51 and the Logic,52 they began to provide linguistic mechanisms, they were interested—they weren’t Cartesians, like they were anti-Cartesian in many ways, but the basic idea that permeated the intellectual atmosphere, they worked out methods of—first of all, they studied the vernacular, which was unusual at the time. Descartes did too, he wrote it in French, which was unusual, but the Port Royal logicians and linguists also studied the vernacular, and they noticed puzzles, in fact there were, there was a famous puzzle, called the Rule of Vaugelas,53 which was a descriptive observation about French, and it was very puzzling at the time, and they worked out an analysis of it, which is based on principles that are very similar to ‘extension’ and ‘intension’ in the Fregean modern sense.

Yes, I see.

And they also put a lot of this in the framework of rules of, sort of like generative rules, it’s a little bit anachronistic, but if you look at it, they’re kind of precursors to generative rules.

OK! Right. I think we are rapidly approaching—

Incidentally, as far as Cartesian—in the same book, I went on to carry this for-
ward to —

Let me get to a question quickly which I think is very often misunderstood, namely, when you wrote in the 1950s, we can’t found syntactic analysis on semantics. I think what’s at issue here is simply the use of the word ‘semantics’, which when by semantics you understand how language is used, which was, —

Yeah, but there was already a straightforward point at the time, which has been totally misunderstood.

Yeah!

The standard view at the time, expressed very clearly by Quine, but also in the linguistics and so on, was that you could do an analysis of language just as a system of form, but you had to appeal to meaning at two points: One point was for phonemic distinctness. In order to show that, you know, let’s say, ‘writer’ and ‘rider’ are distinct, you have to appeal to the meanings. So the—Harris actually had a way around this, but the general view was you had to appeal to the meaning. Also to decide whether a sentence was grammatical, you had to appeal to the meaning. And the specific technical argument having to do with this was to show that neither of these would work. If you actually paid attention to the meaning, you couldn’t determine grammatical status or phonemic distinctness on the basis of the meaning, and in fact, there was really no point at which use of meaning, the facts of meaning, entered into determining what the, say the way a transformational rule works—that has nothing to do with the study of meaning! In fact, LSLT, a lot of which is structuralist, it’s basically a study of semantics! It’s a study of, you know, why the sentences mean what they do. Like, why does “John is eager to please” mean something different from “John is easy to please”. That’s essentially a study of semantics. It’s not based on semantic intuitions; the idea is that we have a formal system in our heads, we use it in particular ways, to express thoughts, to talk about the world, in all kinds of ways. And that’s semantics. Now it’s not even very clear if you look closely, that language has a semantics in the technical sense.

Yes, yeah, absolutely. I think we will tease this very interesting and I think decisive question, we’ll tease out the finer points simply in writing —

OK.

Because —

Too late.

— we are obedient human beings and —

[Points to Ted Passon, sound- & cameraman] He’s tired. He wants to go home.

And we don’t destroy the schedule that Bev wisely set up for you.54

No, OK!

54 Bev Stohl, the ever advertent and sometimes stern guardian of Chomsky’s well-being, who also successfully saw to it that the interview would not last longer than the allotted 60 minutes. With 59 minutes and 36 seconds, Noam and I should not rank badly in a list of well-behaved citizens ready to follow reasonable advice and orders.
Thank you very much, Noam, for doing this.

OK.

And I hope to continue the conversation.

What are you going to do with all this stuff?55

Well, at the moment, it looks like it might be rolled into a book. [...] But first of all, I’ll have a lot of work to do on this.

References


Arnauld, Antoine & Pierre Nicole. 1662/2012. Logic, or the Art of Thinking. London: Forgotten Books.


---

55 That is, the previous interviews with Chomsky himself, but also with all the other linguists involved in this series.
Chomsky, Noam. 2010b. The place of displacement and other questions demanding an answer. Interview with Michael Schiffmann, Cambridge, MA, August 5.


Noam Chomsky  
Massachusetts Institute of Technology  
Department of Linguistics and Philosophy  
77 Massachusetts Avenue, Bldg. 32-D808  
Cambridge, MA 02139  
USA  
chomsky@mit.edu

Michael Schiffmann  
University of Mannheim  
English Department  
Schloss, Ehrenhof West 266  
68131 Mannheim  
Germany  
mikschiff@t-online.de