ONE SMALL HEAD--SOME REMARKS ON THE USE OF 'MODEL' IN LINGUISTICS

BY

YOR ICK W ILKS

DECEMBER, 1971

COMPUTER SCIENCE DEPARTMENT

STANFORD UNIVERSITY

STANFORD ARTIFICIAL INTELLIGENCE PROJECT   DECEMBER 1971
MEMO NO. AIM-157

COMPUTER SCIENCE DEPARTMENT
REPORT NO. 247

ONE SMALL HEAD'.----
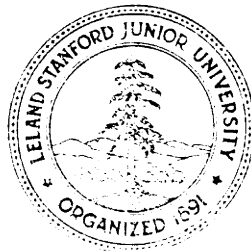SOME REMARKS ON THE USE OF 'MODEL' IN LINGUISTICS

by

Yorick Wilks

ABSTRACT*:   I argue that the present situation in formal linguistics,
             where much new work is presented as being a 'model of the
             brain", or of "human language behavior", is an undesirable
             one. My reason for this judgement is not the conservative
             (Braithwaitian) one that the entities in question are not
             really models but theories. It is rather that they are
             calledmodels because they cannot be theories of the brain
             at the present stage of brain research, and hence that the
             use of "model" in this context is not so much aspirational
             as resigned about our total ignorance of how the brain
             stores and processes linguistic information. The reason
             such explanatory entities cannot be theories is that this
             ignorance precludes any "semantic ascent" up the theory;
             i.e., interpreting the items of the theory in terms of
             observables. And the brain items, whatever they may be,
             are not, as Chomsky has sometimes claimed, in the same
             position as the "occult entities" of Physics like
             Gravitation; for the brain items are not theoretically
             unreachable, merely unreached.

ABSTRACT (Continued)

I then examine two possible alternate views of what
linguistic theories should be proffered as theories of:
theories of sets of sentences, and theories of a particular
class of algorithms. I argue for a form of the latter
view, and that its acceptance would also have the effect
of making Computational Linguistics a central part of
Linguistics, rather than the poor relation it is now.

I examine a distinction among "linguistic models" proposed
recently by Mey, who was also arguing for the **self-**
sufficiency of Computational Linguistics, though as a
"theory of, performance". I arguethat his distinction
is a bad one , partly for the reasons developed above and
partly because he attempts to tie it to Chomsky's in-
scrutable competence-performance distinction. I conclude
that the independence and self-sufficiency of Computational
Linguistics are better supported by the arguments of this
paper.

'And still they gazed, and still the wonder grew,

That one small head could carry all he knew".

Goldsmith's rustics were quite right about the village schoolmaster,
of course, well in advance of their time and, apparently, of Goldsmith.
But perhaps the time has come for less of such gazing, by linguists in
particular, and more attention on their proper business. I am not
suggesting that formal linguistics* has a single proper task, but I am
sure, for reasons I shall try to make clear, that the present situation,
where almost every piece of work in that field is proposed as a new "model
of the human brain or behavior", is an undesirable one. It is not hard
to see in a sympathetic way how linguistics got into that situation. For
a good while there have been serious suspicions, not all voiced from
outside the subject itself, about the other principal explanation of what
linguists were up to, namely providing structural descriptions for sentences.
For it is not at all easy to be clear about the status of conclusions of
the form "X1 is a correct structural description for x, but X2 is not".
Nor has it been merely lack of the appropriate training that has impeded
the understanding of non-linguists, for the experts themselves seemed
to have no way of deciding the truth of such statements in a manner
consistent with the normal standards of rational argument. The pre-
sentation of linguistic work therefore, as being ultimately no less than a
"brainmodel", was a natural, and worthier, alternative to a final
justification in terms of the attachment to sentences of questionable
descriptions.

---

* When I speak generally of linguistics in this paper, it will be clear
  that I am referring to recent developments in the subject and not to
  its traditional comparative and classificatory concerns.

I shall argue here, though, that the present widespread use of "model" in linguistics is unfortunate, above all because it indicates a certain resignation about our almost total ignorance of how the brain actually works. Moreover, I think this situation obscures the proper importance of computational linguistics "(CL), which is capable of providing another, more defensible, justification of the aims of formal linguistics at this stage of neurophysiological research. In due course I shall examine some recent remarks about models by Mey [8] who also seeks to defend the independent position of CL. I have chosen his remarks, rather than other easily available and yet more startling remarks about models by non-computational linguists, only because I agree largely with-what he argues for.

Ten years ago, Chao [3] surveyed the usage of "model" and Suppes [11] has carried out a more rigorous and contemporary study. Both adopted a Websterian, or what might better be called the hundred flowers approach, to the diverse uses of the word in research of that time. They would both, I think, have accepted Mey's opening remark:

"An important notion in the behavioral sciences is that of a model as a set of hypotheses and empirical assumptions leading to certain testable conclusions called predications (on this cf. Braithwaite......"

Now, of course, that is precisely the kind of entity that Braithwaite wrote should be called not a model but a theory, though he did admit that confusion need not necessarily result if "model" is used in this way. [2]

Admittedly Braithwaite's is a conservative view of how "model" should be used. He has tried to assimilate its use in empirical science to its use in mathematics, where it is used to mean a second interpretation of a calculus ielding an understood branch of the subject. The fact that

2

a model, in this sense, exists shows that the first interpretation of the calculus (which is the theory in question, the one being "modelled") is a consistent interpretation. Or, in **Tarski's** words, "a possible realization in which all valid sentences of a theory-*are satisfied is called a model of the theory". Let us call this standard view of **Tarski's** MATH.

Braithwaite's view has been widely discussed, and criticized on the ground that it puts its emphasis on the calculus, and the theory as an interpretation of the calculus, in a way that is untrue to the actual psychological processes of scientists. Opponents of that view argue (Hesse, [7] Achinstein [1]) that the model comes first and that working scientists **import** features metaphorically or analogically from their chosen model into the theory under construction. It may subsequently turn out, as Braithwaite says, that model and theory can be shown to be interpretations of a single calculus, but that is all formal tidying up, such opponents would say, after the real work is over.

However, this difference of views, which I shall call **BRAITH** and SIMPLESCI respectively, is more a difference of emphasis than might appear. Braithwaite, for example, has discussed how within his scheme of things, one can talk of a "modellist" (SIMPLESCI) moving from model to theory by disinterpreting his model's calculus in order to reinterpret it in the terms of the theory paper. Braithwaite contrasts this with his own "contextualist" view that the theory is an interpretation of an originally underlined calculus, with the (**BRAITH**) model entering the picture only subsequently.

For my purpose, though, it is important to emphasize Braithwaite's point that both the BRAITH and SIMPLESCI views of models envisage the

<u>theoretical terms</u> of the theory gaining their interpretations from the bottommost, empirical, level of the theory upwards. Braithwaite refers to this process as a "semantic ascent".*

Mey is of course correct when he says that the "model" is used in a sense different from these three in the behavioral sciences, and in linguistics in particular. The interesting question is, why is it used differently and is there any need to do so, unless something previously unclear is made clear. Let us refer by **MEY** to the view quoted at the beginning; namely that a model is a set of hypotheses, etc., leading to testable conclusions called predictions. That is to say that a model (**MEY**) is what is otherwise called a theory.

Let us now imagine a linguistic (MEY) model. It doesn't much matter what it is, but presumably it will produce word strings of some sort. Those who believe in the overall importance of syntax will want them to be "grammatically correct" strings in one sense; others will want them to be meaningful strings. And that need not just mean sentence strings, in the conventional typesetter's sense, but could mean utterances of any length, including dialogue, that were meaningful and coherent. The distinction
· between "grammatical" and "meaningful" need not disturb us at this point.

After all this clearing of the ground, I want to offer a suggestion as to why the entity under discussion is called a model (MEY) of sanething, rather than a theory of it. The something in question is, of course, the human language apparatus, or part-brain if that is preferred. In the first line of his paper Mey says of CL, as is often said of linguistics in general, that it has to do with human behavior. But that is really a separate matter

---

* My paper does not assume **BRAITH**, but I take it **that** any view of models **requires** some such process, not at present possible in the case of the brain.

4

from the overall point of view of a theory of language production, since
it is agreed by all parties that human language is produced in actual fact
by the human brain and associated organs, hence anything that is to be a
theory of human language behavior must, _ambulando,_ be a theory of some
part of the brain.

My point is concerned with the process of interpretation of theoretical
terms that I referred to earlier: the terms of a theory are given meaning
by reference to lower levels of the theory, that is to say that empirical
base. At the highest levels of a scientific theory there may be entities
with no direct interpretation in terms of the observational base. These,
like "neutrinos" in quantum physics, are sometimes referred to as _occult_
entities, and philosophers go about asking of them "do neutrinos really
exist"? Nonetheless such entities usually have a firm place in a theory
provided they occur only at the topmost levels: in other words, provided
that the process of interpretation can get some reasonable way up from the
observational base.

But in the case of a linguistic theory, proffered as a (MEY) model of
the brain, the situation is quite different from the one I have just
described. At the bottom level, as it were, we can observe what people
actually say and write; or, if one prefers, what they ought to say and write,
it-doesn't matter which for the moment. But, in the present state of
neurophysiological investigations, the matter ends there. There is no
possibility of interpreting further, of identifying any item of structure
in the brain corresponding to any item or structure, at any "level", of the
linguistic theory. And that situation is quite different from that of the
empirical "unreachability", as it were, of neutrinos, for the brain items

5

or structures are not so much unreachable as unreached.

It is this point, I think, that Chomsky missed when he compared [6] the role of unreachable occult entities in linguistics, grammars innate in the mind* in this case, with the positive role of occult entities like Gravitation in Newton's theory. Gravitation features as a topmost item of a theory that admits a paradigmatic semantic ascent from an observational base. But in the case of the brain-there is as yet no agreement at all on how the brain stores and processes information of the type under discussion, and there can be no question of a semantic ascent up a linguistic theory of the brain until that is known.

I think this point explains the MEY use of "model": linguistics cannot provide theories of the brain, or human language production, so what it does provide is called a model . The MEY use expresses an implicit resignation. On the other hand, this usage does undoubtedly express an aspirational, SIMPLESCI, element as well: in that linguistics could, in principle, offer helpful suggestions to brain investigators as to what to look for, though I know of no evidence that such suggestions are being accepted.

Mey is correct then in recognizing, albeit implicitly, that linguistics cannot at present offer full blooded <u>theories</u>** of human

---

* No one should be misled at this point by the fact that Chomsky speaks of "mind" rather than "brain" in the source referred to. In so far as he is speaking particularly in the traditional "mind-mode", his arguments have, I think, all been effectively dealt with in such writing as Putnam (9). I therefore take him to be making remarks about the brain. At other times, of course, Chomsky writes as if such grammars are not occult entities but are actually physically present in the brain.

** There are of course other objections to any apparently deterministic theory of the human brain, language behavior, or whatever; objections well known to any reader of Wittgenstein. But those would only arise when such a theory had actually been produced, and we need not concern ourselves with them here.

language behavior with semantic ascent. However, to go from that to endorse a resigned, and diminished, use of "model" seems to me unfortunate, confusing, and moreover inconsistent with his argument for the independence of computational linguistics as a subject.

But if linguistics cannot provide structures capable of being interpreted as theories of the brain, and so of human behavior, and if, also, the depleted sense of "model" is less than adequate to cover these would-be theories, is there then any other alternative? It surely cannot be enough for formal linguistics to go into an academic hibernation to await a breakthrough in the description of the brain itself. If linguistics is to offer theories, what are they to be described as theories of? Two obvious alternatives present themselves: firstly, that what linguistics provides are ultimately theories of sets of sentences. Secondly, that what 'it provides are theories of a particular class of algorithms.

Even if "sentences" is taken in a wide sense, so as to include whole discourses of any reasonable length, there seem strong and traditional objections to the first of these suggestions. For the proposal may sound like no more than a resurrection of the form of logical empiricism in the philosophy of science most closely associated with Neurath. [9] **For** Neurath, a theory was no more or less than a production system for the basic sentences, or **Protokolsätze**, of a science. Beyond that the theory was wholly dispensable, and there was no place in his views for models of any sort, or "semantic ascents" up the levels of theory.

There are well known objections to such a view of theories in general. From the standpoint of the argument expressed here, the view is unacceptable because a linguistic theory that was merely a theory of a set of sentences, with no add'l qualification, would, on the Neurathian view of theories,

7

also be a theory of anything producing such a set of sentences, and among these things would be human beings and their brains. That is the view of linguistic theory, of course, which makes the easy and almost imperceptible shifts of many generative grammarians, between talking of theories of sentences and theories of human brains and behavior, most plausible and acceptable. However, it is a general view of theories which, if they thought about it, most of them would wholeheartedly reject: Chomsky himself, for example, has argued many times against any such empiricist view of theories.

Chomsky himself makes these transitions frequently, though he is by no means a consistent user of "model" to mean "theory" in this context in the way I argued against earlier. He frequently writes of theories, though in a number of different ways:

"There is a certain irreducible vagueness in describing a formalized grammar as a theory of the linguistic intuition of the native speaker". (Chomsky 5, p. 533)

Chomsky is arguing for such theories here, of course, and this is a formulation of his position apparently different from any of the views of the role of theories mentioned so far: the sentence view, or the grammar-in-the-brain-or-mind views.* However, if we ignore the limitation to grammar in any narrow sense, this statement reduces to something very like the linguistic-theories-are-of-sentences view under discussion, at least if the intuitions in question are restricted to intuitions as to what are and are not sentences.

---

* I have argued elsewhere and in detail [14] against the way in which Chomsky makes these transitions, and also that these intuitions that justify anv particular set of sentences cannot, whatever they are, be syntactic ones in any serious sense. But that disagreement need not affect the point under discussion.

8

I would myself suggest a version of the second of the above views,
namely that we view linguistic theories as the production of particular
sets of sentences by programs or algorithms. There is an implicit
restriction included there, naturally enough, to non-trivial methods that
would exclude the printing out of any prestored list of sentences. That
formulation may sound like no more than an analytic definition of the phrase
"computational linguistics", and indeed, as Professor John Wisdom has so
often pointed out, philosophical proposals are usually no more than the
announcement of a platitude. But in the current state of the use of "model"
and "theory" in linguistics, any <u>single</u> way of speaking of theories would
be an advantage **if it replaced** the current Babel in a generally acceptable
manner. Most importantly, and here I think Mey might agree with me, the
proposed view of theories in linguistics would make CL the foundational part
of formal linguistics, and not the poor relation it is treated as at present.
Yet, if, as Chomsky has always argued, linguistics is to be more than the
mere classification and comparison it used once to be, then I do not see how
generative grammarians can resist some such view, of what linguistic theories
are theories of, as the one proposed here.

On this view, the items of a linguistic theory could, without too
much difficulty, be identified with subparts of the algorithm, in a way that
cannot be done for the brain. More importantly, this view could be related
in a coherent fashion to current notions of theory and model, and in that
sense would have an obvious advantage over the loose talk of "psychological
modelling" with which contemporary linguistics is so beset. For **example,**
it would be possible for such a theory of CL to have a model (BRAITH and
MATH), in the sense of an area of logic or mathematics with suitably related

9

properties. These models almost certainly exist for a number of CL

theories: those using phrase-structure algorithms, for example. Again,

there is no reason why what people say about their language structure,

and what facts psychological experiment-can elicit about the associations

between speech items, should not serve as suggestive models (SIMPLESCI)

for proper theories of CL. Linguists who are wondering if they read that

last sentence the right way round, need not read it again, they did*.

It may be objected at this point that such a view is too <u>particular.</u>
Given the flourishing state of Automata Theory proper and the **theory** of

algorithms, whether viewed as a part of mathematics, logic, or mathematical

linguistics, it **is~as** absurd to suggest this view of CL as to seek to

propagate the "Chemistry of the Apple" as an independent subject. However,

there need be no conflict here, and on the view under discussion it would

be quite reasonable to conceive of CL within either mathematical linguistics

or automata theory as their implemented aspect, one which might be expected

<u>ipso facto</u> to be less mathematically interesting than the general theory

of algorithms or the theory of abstract machines. There can be no objection

in principle, though, to a CL theory being a theory of algorithms, on the

grounds that the algorithm might have been described in some other way. At

least, not if the objector is a linguist who does want "psychological **modelling**"

and theories of the brain, for he would hardly take it as an objection

to some future theory of language production in the brain that the area of

the brain in question might just as easily have been used to process, say,

visual data.

---

*Those who, like Hesse [7], adopt an "interaction" view of the role of
  (SIMPLESCI) models **would** say that this possibility was to be expected.

There are two **ancilliary** arguments which, I think, justify the introduction of the notion of an algorithm into the definition of a theory of formal linguistics.

Firstly, it is a fact of academic observation, as I mentioned earlier, that the descriptions which linguists provide for utterances are disputable, what one might call **undecidably** so. The production, or non-production, of strings (or analysis, or non-analysis, of course) by algorithm provides a nondisputable justification for whatever linguistic classification and description had been initially imposed and programmed. To put an old and well-labored point briefly: classification, in linguistics at least, requires sane purpose, or something one wants to do, and CL can provide it.

It is not usually necessary to operate a <u>logical system</u> very far in order to see whether or not it produces the set of strings that are in question, the theorems, for that can usually be seen by inspection. But the rules of linguistics are generally so much more numerous and complicated that inspection is not usually sufficient. Furthermore, inspection in such cases is prey to all the well known weaknesses of investigators for looking for what supports their case and ignoring what does not. If the strings are produced by algorithm, possibly out of a machine, it is more difficult to select unconsciously in that way.

Let us now take a warning look at the distinction among models (MEY) that Mey actually proposes. The (BRAITH) theory, or (MEY) model, Mey proposes to call a <u>descriptive model</u> (MEYD). He cautions us that "it need not be (and should not be) considered a faithful reproduction of reality, in the sense that to each part of the model there corresponds, by some kind of isomorphic mapping a particular chunk of 'real' life. In other words, **this**

descriptive kind of model does not attempt to imitate the behavior of the descriptum". (p.2).

The last sentence might leave one asking, well if it doesn't do that, what does it do that deserves interest and attention? There is also an ambiguity about notion of "mapping" here. It might seem that by 'mapping" Mey refers to the interpretation of model (MEYD) items at different levels by brain items. But he goes straight on to discuss the non-equivalence of behavior which suggests that assumption is wrong, and that he means only that the model (MEYD) need not even give output like human behavior.

He goes on, "The other kind of model I propose to call the simulative one . . . . . .a conscious effort to picture, point by point, the activities that we want to describe" (ibid. pp. 2-3). The elucidation of the distinction, between MEYD and what I shall call MEYS, is wholly in terms of that philo-sophical monster, Chomsky's competence-performance distinction. For example, MEYD models are said to be like Chomsky's competence models, yet he writes of MEYD's: "The model that is a grammar does not attempt to explain linguistic activity on the part of the speaker or hearer by appealing to direct similar-ities between that activity and the rules of the grammar. Rather, the activity of the speaker (his performance) is explained by pointing to the fact that the rules give exactly the same result (if they are correct that is) as does the performance of the speaker-hearer." (p.4)

But there is serious trouble here. If these two entities, the human and the MEYD, give the same result then, as I have pointed out at some length, the one does not explain the other in the sense that an interpreted theory explains what it is a theory of. Mey writes in the last quoted passage as if there are other similarities, between grammar and human, other than output identity. But what can they be? Moreover, dubious as Chomsky's

distinction is, I am not sure that Mey has got it right, for which he can be forgiven, of course. For it seems odd to identify **MEYD** with Chomskyan comparison between the 'outputs' of grammars and humans, since that is surely within what Chomsky would call 'performance'.

Again, the last quoted passage makes clear that **MEYS's** describe by definition, just as do **MEYD's**. According to the definition given, the distinguishing feature of **MEY's** is that they picture 'point by point' the human language activity. But, as I have argued at some length above in connection with the general notions of model and theory, that is just what they cannot conceivably do, at least not at the moment while there is no hint available as to what the 'points' to be pictured are. In the case of human beings as of machines, output is output is output, so what distinction can Mey offer between **MEYD's** and **MEYS's**, since ultimately all they both have to 'model' is human output?

MEY models won't do, partly because one can do better with the notion of a theory of CL, and partly because the distinction between MEYD and MEYS is tied to the inscrutable competence-performance distinction. Chomsky means different things by that at different times: if he is attacked on one version he shifts to another (see 10 and 14). As Paul Garvin put the matter some while ago "with the new linguistics, came a new style of argument". To take just Chomsky's version of competence that Mey begins with (p. 2) as "a speaker's knowledge of a language", which Chomsky takes essentially to include a grammar , yet as should be widely known by now, the majority of the world's competent speakers probably do not even know their language has a grammar.

The main point that Chomsky has tried to express by means of that

distinction is, I think, that the behavior of all actual things differs significantly from theoretical predictions and idealizations. Actual billiard balls collect dust as they roll and slow down, so their experimental performance never quite lives up to their theoretical mathematical competence. People as speakers are no different in this respect from other things in the natural world: they insert **'ums'** into otherwise perfect sentences. But this is true in **every field,** and deserves no special terminological recognition in the case of human speech or writing.

In his paper Mey is arguing for worthwhile things, and in particular that linguistics, by which he means CL, should concern itself more with meaning and less with grammar, and that it should concentrate on the acceptance and interpretation of utterances rather than their **acceptance-or-rejection** offhand. He argues that the latter will require new kinds of theories in CL. I agree and have tried in (13) to suggest what their general form might be. His mistake, I think, is to try and make these valuable points with the aid of an ill-thought out distinction between two kinds of 'models'. I say ill-thought out advisedly and for two reasons: first, as I argued in detail, 'model' is best kept for other and more conventional uses, and CL would benefit more from a suggested extension of the term 'theory'. Secondly, because Mey thinks that whatever it is he has **to say,** it must have something to do with what Chomsky meant when he at **various** times tried to distinguish competence and performance, and in particular that CL can find an acceptable theoretical niche by being the long-awaited 'theory of performance'.

If my general argument in this paper is correct there <u>can</u> only be <u>'theories of performance'</u>, and for that CL is a foundation stone and in no need of a niche in **linguistics.**

REFERENCES

[1] P. Achinstein, "Theoretical Models", <u>Brit. Jl, Phil. of Sci.</u>, 1965.

[2] R. B. Braithwaite, "Models in the Empirical Sciences", in (eds.) Nagel, Suppes and **Tarski,** <u>Proc. 1960 International Congress on Logic, Methodology and Philosophy of Science.</u> Stanford, 1962.

[3] Y. R. Chao, "Models in Linguistics and Models in General" as (1).

[4] N. Chomsky, "Three Models for the Description of Language", <u>IRE Transactions in Information Theory</u>. Vol. II-2 (1956).

[5] N. Chomsky, "Explanatory Models in Linguistics" as (1).

[6] **N.** Chomsky, <u>Language and Mind</u>, New York, 1969.

[7] M. Hesse, "The Explanatory Function of Metaphor" in <u>**Proc. 1964** International Congress on Logic, Methodology and Philosophy of Science,</u> Amsterdam, 1965.

[8] **J. Mey,** "Toward a Theory of Computational Linguistics", presented at the Annual Meeting of the <u>Association for Computational Linguistics,</u> Columbus, Ohio, 1970.

[9] 0. Neurath, "Protocol Sentences" in (ed) A. J. Ayer, <u>Logical Positivism</u>, Glencoe, Illinois, 1959.

[10] H. Putnam, "The 'Innateness Hypothesis' and Explanatory Models in Linguistics", <u>Synthese</u> 17, 1967.

[11] P. Suppes, "A Comparison of the Meaning and Uses of 'Models' in Mathematics and the Empirical Science", <u>Synthese</u> 12, 1960.

[12] Y. Wilks, "On-line Semantic Analysis of English Texts", <u>Mechanical Translation and Computational Linguistics</u>, 1968.

[13] Y. Wilks, "Decidability and Natural Language", <u>Mind</u>, 1971.

[14] Y. Wilks, <u>Grammar, Meaning and the Machine Analysis of Language,</u> London, 1972.