June 17 2009
Keio University
Main Handout

Generative Grammar as an Empirical Science: its goal and how it can be pursued—with some historical reflections

Hajime Hoji
University of Southern California

This talk is a somewhat non-technical elaboration of the thesis put forth in my book manuscript *A Foundation of Generative Grammar as an Empirical Science* (henceforth Hoji 2009) and its general significance and implications. The issues to be addressed include:

- The goal of generative grammar (as an empirical science)
- What has to be ensured if we are to pursue the goal with the general scientific method that seems to be accepted fairly widely: Guess-Compute Consequences-Compare, Experiment, as Richard Feynman puts it
- How we might try to (i.e., how I think we can) make generative grammar an empirical science in line with the general scientific method alluded to above
- Some implications of the proposal

The talk will also address:

- How the field of generative grammar at large has proceeded in relation to its articulation of its methodology
- Some of its negative consequences
- How the proposed methodology in Hoji 2009 can help us avoid such problems

Because of the time constraints, I regret that I cannot provide (almost) any concrete illustration of the proposal, including how specific hypotheses get revised or rejected in accordance with its methodological proposal. The separate handout is meant to give the reader *some* idea about the empirical discussion (as well as what is meant by a few of the notions put forth) in Hoji 2009.

1. *Generative grammar* as the study of the language faculty

- What is the goal of generative grammar?

Chomsky's remarks in *Third Texas Conference on Problems of Linguistic Analysis in English* May 9-12, 1958, published in 1962, seem to point directly to what he had in mind at least around 1958, in my view, more directly than what we find in his writings in the 1950s and 1960s and the subsequent years. (The emphases in (1) and (2) are by HH.)

(1) (p. 167)
Hill: If I took some of your statements literally, I would say that you are not studying language at all, but some form of psychology, the intuitions of native speakers.
Chomsky: *That is studying language*.
Long: I agree with Chomsky and Harris here. Language goes on in the brain, not merely in the throat.
Chomsky: How language fits into the throat is a matter which is quite interesting. I claim, however, that study of *the native speaker's reactions is what all linguists are studying.*[1]

(2) (p. 168)
Chomsky: I don't think such a test eliminates intuition; I think we want our tests to converge on intuition. If you want to eliminate intuition, then I think my absurd procedure is perfectly satisfactory.
Hill: Linguistic intuition is itself a system, almost a complete grammar. If it is good enough, why bother with any other grammar?

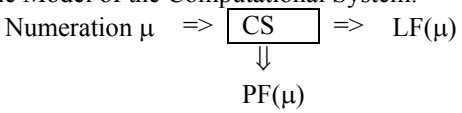
[1] There seems to be a typo here. But I am reproducing what is in the volume.

Chomsky: *Because I am interested in explaining intuition. If you cannot accept this as the purpose of linguistic study, I am lost. I would like to get a theory which will predict intuitions.*

- Minimally, the language faculty must relate "sounds" (and signs in a sign language) and "meanings."

A fundamental hypothesis in *generative grammar* is the existence of the Computational System at the center of the language faculty. Since Chomsky 1993, the Computational System is understood in generative research to be an algorithm whose input is a set of items taken from the mental Lexicon of the speaker of a language and whose output is a pair of mental representations—one underlying 'sounds/signs' and the other 'meaning'. Following the common practice in the generative tradition since the mid 1970s, let us call the former a *PF* (representation) and the latter an *LF* (representation). The LF and the PF representations in (3) are thus meant to be abstract representations that underlie a sequence of sounds and its 'interpretation', respectively.

(3) The Model of the Computational System:



Numeration: a set of items taken from the mental Lexicon
LF(μ): an LF representation based on μ
PF(μ): a PF representation based on μ

The hypotheses in (4) are among the *hard core* (in the terms of Lakatos' work) I adopt.

(4) a. The Computational System exists at the center of the language faculty.
b. The general organization of the Computational System is as indicated in (3).[2]

I adopt a very small number of additional hypotheses in the *hard core*, such as the nature of the operation in the Computational System, *Merge* (external and internal). Any hypotheses other than those belonging to the *hard core* are subject to empirical tests, modification, and abandonment. They correspond to Lakatos' *auxiliary hypotheses*.

- Our hypotheses about the Computational System are thus meant to be about *what underlies* the language users' intuitions about the relation between "sounds" (and signs in a sign language) and "meanings" and the methodological proposal put forth in Hoji 2009 is how to make testable predictions and how to evaluate the significance of the results of our experiments.

The above characterization of the goal of *generative grammar* is very much in line with Chomsky 1965.

(5) Chomsky 1965
a. [L]inguistic theory is mentalistic, since it is concerned with discovering a mental reality underlying behavior. (p. 4)
b. Mentalistic linguistics is simply theoretical linguistics that uses performance as data (along with other data, for example, the data provided by introspection) for determination of competence, the latter being taken as the primary object of its investigation. (p. 193 in a note appended to (5a))

One important caveat: Hypotheses about the properties of LF (and PF) can be understood as being about the Computational System in a broad sense. The Computational System as schematized in (3) is the narrow sense of the Computational System.

2. What has to be ensured

In Hoji 2009, I explore some consequences of adopting the following general scientific method (as Richard Feynman puts it), for research concerned with the properties of the language faculty.

(6) The general scientific method:
Guess — Computing Consequence — Compare, Experiment

[2] Included in the hypothesis in (3) is the existence of a mental lexicon.

YouTube video: http://www.youtube.com/watch?v=knDXAr4ltMA
Feynman Part 1 (about 3' 15" from the beginning of the tape) (See (52b) in Appendix II below.)

"I am going to discuss how we look for a new law. In general, we look for a new law by the following process. First we guess it [Writing on the blackboard "Guess."]. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. [Writing on the blackboard "Computing Consequences."] Then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works. [Writing on the blackboard "Compare, Experiment."] Compare it directly to observations to see if it works. If it disagrees with experiment, it's wrong. In that simple statement is the key to science. It doesn't make any difference how beautiful your guess is, how smart you are, who made the guess, or what his name is. If it disagrees with the experiment, it's wrong. That's all there is to it."[3]

Feynman continues the above passage by adding "obvious remarks," stating, "When I say if it disagrees with experiment it is wrong, I mean after the experiment has been checked, the calculations have been checked, and the thing has been rubbed back and forth a few times to make sure that the consequences are logical consequences from the guess, and that in fact it disagrees with a very carefully checked experiment."

More in particular, in Hoji 2009 I propose that the language faculty can be studied with the general scientific method as schematized in (6) and make concrete suggestions as to how.

In order to proceed with our investigation of the language faculty with the general scientific method in (6), we must ensure the following, at least to a minimally satisfactory degree.

(7) a. It is possible to compute the consequences of the "guess."
b. It is possible to determine whether or not the consequences of the "guess" agree with the observations and/or the experimental results.

(7a) and (7b) can in turn be ensured only if (8) and (9) hold, respectively.

(8) The "guess" is part of, or is related to, a larger deductive system.

(9) The consequences of the "guess" are related, ultimately, to something 'observable/measurable'.

Furthermore, we should ensure (10) as best as we can.

(10) Disagreement between the consequences of the "guess" on the one hand and the observations and/or the experimental results on the other could lead us to learn something about the language faculty.

3. On the need of methodology

3.1. Chomsky on methodology

In answering a question about [his] "method of investigation", Chomsky (1980[4]: 190) states as in (11).

(11) As for my own methods of investigation, I do not really have any. The only method of investigation is to look hard at a serious problem and try to get some ideas as to what might be the explanation for it, meanwhile keeping an open mind about all sorts of other possibilities. Well, that is not a method. It is just being reasonable, and so far as I know, that is the only way to deal with any problem, whether it is a problem in your work as a quantum physicist or whatever.
There are certain fields like psychology where people do carry out extensive study of methods of investigation. There are other fields like physics where you do not study methods of investigation. So at MIT the physics department does not have a course in experimental methods, but many psychology departments spend a lot of time on what they call methodology. Well, there is a lesson there, but I won't draw it. (Chomsky 1988: 190)

[3] The passage that corresponds to the above is in: Feynman, Richard. 1994. *The Character of Physical Law*. New York: The Modern Library. (p. 150) (The book was originally published in hardcover by BBC in 1965 and in paperback by MIT Press in 1967.)

[4] Chomsky, Noam. 1988. *Language and Problems of Knowledge: The Managua Lectures*, MIT Press.

The lesson that Chomsky did not draw seems to be something like (12).

(12) If you are doing real science, you don't talk about methods. After all, they don't do that in physics, perhaps the most advanced area of science.

Well, there is a lesson to draw here, and I would like to draw it.

(13) In physics, at least in the field as a whole, there has been a long, in fact very long, tradition of empirically testing hypotheses as a fairly reliable means to determine the plausibility of the hypotheses. Even so-called theoretical physics could not have a meaningful existence if it were totally unrelated to empirical observations/generalizations.

--For hypothesis forming, there is no method. We know that; well, at least, most people seem to agree on that. But for hypothesis testing, the situation is totally different. I do not mean to endorse the practice in psychology that Chomsky seems to have alluded to in (11). To the extent that much of psychology is concerned with behavior rather than the nature/representation/mechanisms of the mind, something is seriously wrong with their emphasis on "methods," I would agree. But that does not mean that we need not be concerned with methods of *testing* our hypotheses. This is related to the nature of data in the empirical inquiry that we are engaged in.

- A crucial question: *What should count as data in research that is concerned with the properties of the language faculty*, and more narrowly, with the properties of the Computational System that is hypothesized to be at the center of the language faculty?

One might be tempted to infer from (11) that Chomsky refuses to draw a line between testable hypotheses and not-testable hypotheses (i.e., between scientific hypotheses and metaphysical (or pseudo-scientific) hypotheses) noting that a serious astrologer perhaps did exactly what is suggested in (11).[5] While others might also wonder whether the view expressed in (11) can be reasonably regarded as Chomsky's consistent position, such does indeed seem to be the case, given that we do not seem to find Chomsky's remarks on methodology that is contrary to (11) and that we do find Chomsky's remarks, directly or indirectly, that are consistent with (11). Schütze 1996[6], for example, states as in (14).[7]

(14) Chomsky (personal communication) believes that research practice in linguistics ought to follow that in the natural sciences, where (in contrast to the social sciences) "almost no one devotes attention to 'methodology'." Obviously, I disagree.
(Schütze 1996: 210, footnote 1)

3.2. Chomsky on history and philosophy of science

Consider also (15), taken from Chomsky 1979.[8]

(15) I should also mention work on history and philosophy of science, which has begun to furnish a richer and more exact understanding of the manner in which ideas develop and take root in the natural sciences. This work—for example, that of Thomas Kuhn or Imre Lakatos—has gone well beyond the often artificial models of verification and falsification, which were prevalent for a long time and which exercised a dubious influence on the "soft sciences," as the latter did not rest on the foundations of a healthy intellectual tradition that could guide their development. It is useful, in my opinion, for people

---

[5] 湯川秀樹 理論物理学の輪郭 in「目に見えないもの」 pp. 12-13 contains remarks that seem relevant here. (first published in 1946, the page references are to the 1976 講談社学術文庫.)

[6] Schütze, Carson. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*, University of Chicago Press.

[7] (14) is given as a footnote appended to the text remark in (i).

(i) ... I would echo Greenbaum's (1977c) recommendation that every linguistics department should offer a course in experimental linguistics. In addition to reasons internal to our own field, this would give a student a leg up in joining the blossoming interdisciplinary enterprise of cognitive science. It would also seem to be a natural outgrowth of Chomsky's own suggestion that linguistics be viewed as a branch of cognitive psychology. Somehow, the focus on cognitive issues has not yet been accompanied by adoption of the scientific standards and concern with methodology of that discipline. (Schütze 1996: 210)

[8] Chomsky, Noam. 1979. *Language and Responsibility: Based on conversations with Mitsou Ronat*, Panthen Books, New York.

working in these fields to become familiar with ways in which the natural sciences have been able to progress; in particular, to recognize how, at critical moments of their development, they have been guided by radical idealization, a concern for depth of insight and explanatory power rather than by a concern to accommodate "all the facts"—a notion that approaches meaninglessness—even at times disregarding apparent counterexamples in the hope (which at times has proven justified only after many years or even centuries) that subsequent insights would explain them. These are useful lessons that have been obscured in much of the discussion about epistemology and the philosophy of science. (Chomsky 1979: 73)

In order to be able to evaluate what seems to be intended by the above, we need to understand what is meant by each of (16).

(16) a. the often artificial models of verification and falsification
b. the "soft sciences"

We should also address (17).

(17) what counts as a fact

My own research methodology has been to proceed with the Reinhartian heuristic (see below), and to do so in an empirically testable manner it is imperative that we commit ourselves to *pf-LF correspondences*, as discussed and illustrated in some depth in Hoji 2009.

The two most general hypotheses put forth in Reinhart 1983 are (18) and (19).

(18) Reinhart 1983: 25, (19):
Sentence-level semantic interpretation rules may operate on two given nodes A and B only if one of these nodes is in the domain of the other (i.e., A is in the domain of B, or B is in the domain of A, or both).

(19) Reinhart 1983: 26, (21):
If a rule assigns node A some kind of prominence over node B, A must be a D-head of the domain which contains B.

We can restate (18) and (19) as in (20) and (21), respectively, taking "sentence-level semantic interpretation rules" as "CS-based rules or conditions that contribute to or regulate interpretive possibilities," which seems to be a reasonable interpretation, given the discussion in Reinhart 1983.

(20) Reinhart 1983: 25, (19), restated:
CS-based rules or conditions that contribute to or regulate interpretive possibilities can involve A and B only if A c-commands B, or B c-commands A, or both.

(21) Reinhart 1983: 26, (21), restated:
If B is *dependent upon* A in terms of how B gets interpreted, B must be c-commanded by A.

In the context of the present discussion, let us take the statements in (20) and (21) as applying at LF.[9] (20) is a hypothesis about the form of a CS-related hypothesis that pertains to an interpretation involving two elements; (21) is an instance of a bridging statement. We can thus understand Reinhart 1983 as containing a heuristic like (22).

(22) The Reinhartian heuristic:
The relation at LF that underlies γ(a, b) mentioned in a bridging statement must be based on a *c-command* relation between LF(a) and LF(b).

It seems to me that what is meant by (16a) is the models by philosophers in which verification and falsification

[9] Although taking (18) and (19) (hence (20) and (21)) as applying at LF is not justified by the textual reading of Reinhart 1983 alone—while the restatement of (18) and (19) as in (20) and (21) can—, that is a reasonable way to understand what is intended in Reinhart 1983 once it is 'translated' into the general framework adopted here regarding the organization of the CS.

of a scientific hypothesis is to be done on the basis of some concrete observational facts, or something like that (that, I understand, is the basic tenet of "logical positivism"). I assume that what is meant by (16b) are inquires beyond the 'natural sciences' (such as social sciences, and perhaps linguistics being included here, as it is practiced in much of the field, (including much of the generative research, I hate to say)).

It might not be an overstatement that how one understands (23), contained in (15), in the context of generative grammar, and more in particular in the context of assessing one's own research, may 'define' one's orientation as a researcher dealing with 'language'.

(23) [A]t critical moments of their development, [the natural sciences] have been guided by radical idealization, a concern for depth of insight and explanatory power rather than by a concern to accommodate "all the facts"—a notion that approaches meaninglessness—even at times disregarding apparent counterexamples in the hope (which at times has proven justified only after many years or even centuries) that subsequent insights would explain them.

One might take (23) to mean that we should not be too concerned with, or impressed by, empirical observations since, after all, empirical observations alone never determine the fate of a scientific theory, anyway.

It may thus be a useful exercise to raise the following questions in relation to the type of research advocated in Hoji 2009, and we might also ask whether the answers to each of the questions would hold only "at critical moments."

(24) a. Is it guided by "radical idealization"?
b. Is it guided by a concern to accommodate "all the facts"?
c. Is it guided by "a concern for depth of insight and explanatory power"?
d. Does it at times disregard apparent counterexamples?

Notice that the content of the crucial notions in (24c) are quite obscure and can be very subjective, as addressed in Postal 2004; chapter 12 though from a somewhat different perspective than what is being pursued here.[10] If someone proposed or endorsed something like (25), that should be reason for concern, in my view; cf. Feynman's remark on the general method in science.

(25) "[A] concern for depth of insight and explanatory power" should guide our research more than a concern for how our predictions are borne out.

As long as we articulate our hypotheses rigorously enough so as to ensure their empirical testability, we should be able to make definite and testable predictions; cf. *severe tests* in Mayo 1996 (*Error and the Growth of Experimental Knowledge*). By carefully designing an experiment to test the predictions, we should be able to tell, fairly objectively, whether or not the predictions are borne out and hope to be able to evaluate the validity of hypotheses about the Computational System reliably. We would face a very different situation if we tried to determine how insightful a given hypothesis or theory is or how much explanatory power it may have, because there does not seem to be any criteria for the relevant determination; see the discussion in the next subsection.

We can also address the following part of what is quoted from Chomsky 1979: 73.

(26) (Cf. (15), which contains this.)
"[E]ven at times disregarding apparent counterexamples in the hope (which at times *has proven justified only after many years or even centuries*) that subsequent insights would explain them." (the emphasis by HH)

We should in particular consider what, in principle, would be required for something like this to happen/take place, i.e., what would be required in order for "apparent counterexamples being disregarded for many years or even centuries and getting explained by subsequent insights, thus justifying the disregarding of them (for many years or even centuries)."

- Do we have such counterexamples in generative grammar over the past half century?

- What *could* be (the form of) such a counterexample in research concerned with the language faculty?
- What would be a necessary requirement for something to be such a recalcitrant counterexample in a given research program?

[10] Postal, Paul. 2004. *Skeptical Linguistic Essays*. Oxford University Press.

Cf. Lakatos' *scientific research programs*.[11]

Two illustration of the hard core, auxiliary hypothesis, progressive and degenerating problemshift:[12]
(i) the anomalous orbit of Uranus and the subsequent discovery of Neptune
(ii) the 'anomalous perihelion precession' of Mercury

3.3. Chomsky on Peirce

Responding to Mitsou Ronat's (27), Chomsky (1979) remarks as in (28).

(27) "Abduction is, I believe, a form of inference which doe not depend solely on a priori principles (like deduction), nor solely on experimental observation (like induction). But that aspect of Peirce is very little known in France."

(28) "Or here in the United States either. Peirce argues that to account for the growth of knowledge, one must assume that "man's mind had a natural adaptation to imagining correct theories of some kinds," some principle of "abduction" which "puts a limit on admissible hypothesis," a kind of "instinct," developed in the course of evolution?" (Chomsky 1979: 71)

If what is intended in (28) were that we must have some intuition about "correct theories," and that we must rely on such intuition in deciding on "correct theories," such a view would amount to a declaration that generative grammar cannot be an empirical science, as far as I can tell. It is possible, if not likely, that Chomsky is comparing, in his remarks above, generative grammar on the one hand and structural linguistics and some behaviorist/empiricist-oriented research activities on the other.[13] I would nevertheless like to think that his remarks above and the brief discussion I have provided above is in fact quite suggestive as to what Chomsky considers (his) generative grammar to be.

Conspicuously missing in Chomsky's remarks above (and remarks here and there in Chomsky 1995: chap. 1, as well) are concerns about how to test predictions that our hypotheses give rise to. I am inclined to think that "depth of insight and explanatory power" can be meaningfully addressed *only if* our research activities are accompanied by concerns with how to test our hypotheses (and the practices in accordance with such concerns).

Chomsky, Noam. 1979 *Language and Responsibility: Based on conversations with Mitsou Ronat*, Panthen Books, New York, 70-71 contains the following.

M.R.: To what degree can your discoveries about language and your definitions of fields of knowledge lead to the emergence of new philosophic questions? To which philosophy do you feel closest?

N.C: In relation to the questions we have just been discussing [which has to do with how we acquire our knowledge in general --HH], the philosopher to whom I feel closest and whom I'm almost paraphrasing is Charles Sanders Peirce. He proposed an interesting outline, very far from complete, of what he called "abduction" … (The "..." part is as in the original. HH)

M.R.: Abduction is, I believe, a form of inference which does not depend solely on a priori principles (like deduction), nor solely on experimental observations (like induction). But that aspect of Peirce is very little known in France.

N.C.: Or here in the United States either. Peirce argues that to account for the growth of knowledge, one must assume that "man's mind had a natural adaptation to imagining correct theories of some kinds," some principle of "abduction" which "puts a limit on admissible hypothesis," a kind of "instinct," developed in the course of evolution.

Remarks by HH:
While the aspects of C. Peirce's concern that interest me most include not only how a hypothesis in science can be formed (I guess by abduction (and also by induction, too, I would suspect, insofar as adduction is possible only if

[11] Hoji 2009: chapter 5 states its methodological proposal in the terms of Lakatos' *scientific research program*.

[12] A question, however, remains as to whether physics is the right place to turn to for our model.

[13] It may be the case that what Chomsky had in mind here was not just the structuralist/behaviorist approach in general but it was directed also toward generative semantics.

there is a generalization of some sort, without which we could not be able to identify a puzzle, to begin with) but also how it should be put to empirical tests (on the basis of deduction (as well as induction, I suspect)), Chomsky, in the above passage, seems to be talking about how we acquire our knowledge in general. It is, however, not clear that Chomsky finds it unnecessary for us to put our (scientific) hypotheses to empirical tests. I would like to think that he does. It might, however, be the case, indeed, that he is not concerned as much as we are with putting our hypotheses to empirical test, and that might not be a particularly unwarranted guess, on the basis of his remarks such as those quoted above as well as on the basis of how he has proceeded with his research especially since the mid 1980s.[14]

3.4. Feynman on methodology

- Richard Feynman's remarks on methodology; cf. Chomsky's remarks on methodology in generative grammar
  - ✧ "Scientific Honesty"

[Taken from Richard Feynman's 1974 Caltech Commencement Address "Cargo Cult Science"; see (54b) in Appendix II below. See also "Feynman on Social Sciences" mentioned in (52a) in Appendix II.]

But there is one feature I notice that is generally missing in cargo cult science. That is the idea that we all hope you have learned in studying science in school—we never say explicitly what this is, but just hope that you catch on by all the examples of scientific investigation. It is interesting, therefore, to bring it out now and speak of it explicitly. It's a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty—a kind of leaning over backwards. For example, if you're doing an experiment, you should report everything that you think might make it invalid—not only what you think is right about it: other causes that could possibly explain your results; and things you thought of that you've eliminated by some other experiment, and how they worked—to make sure the other fellow can tell they have been eliminated.

Details that could throw doubt on your interpretation must be given, if you know them. You must do the best you can—if you know anything at all wrong, or possibly wrong—to explain it. If you make a theory, for example, and advertise it, or put it out, then you must also put down all the facts that disagree with it, as well as those that agree with it. There is also a more subtle problem. When you have put a lot of ideas together to make an elaborate theory, you want to make sure, when explaining what it fits, that those things it fits are not just the things that gave you the idea for the theory; but that the finished theory makes something else come out right, in addition.

In summary, the idea is to give all of the information to help others to judge the value of your contribution; not just the information that leads to judgement in one particular direction or another.

...

We've learned from experience that the truth will come out. Other experimenters will repeat your experiment and find out whether you were wrong or right. Nature's phenomena will agree or they'll disagree with your theory. And, although you may gain some temporary fame and excitement, you will not gain a good reputation as a scientist if you haven't tried to be very careful in this kind of work. And it's this type of integrity, this kind of care not to fool yourself, that is missing to a large extent in much of the research in cargo cult science.

A great deal of their difficulty is, of course, the difficulty of the subject and the inapplicability of the scientific method to the subject. Nevertheless, it should be remarked that this is not the only difficulty. That's why the planes don't land—but they don't land.

- So, after all, according to Feynman, there is methodology in physics.

--One crucial question one might raise is how we can 'measure" scientific integrity/honesty. Suppose one adopted the view that the informant judgments *must converge* on *every single sentence* used in an experiment. One would in

[14] The conspicuous absence of "acceptability judgment," "acceptability and grammaticality," etc. in the Indexes in Chomsky's books since the late 1970s, as compared to his books in the 1950s and 1960s, is also suggestive in this regard; cf. Chomsky 1965 (*Aspects*): chapter 1.

that case most likely feel that the use of informant judgments itself is an act of dishonesty because such convergence just cannot seem to be attained. If one adopts a different view, in line with what is suggested in Hoji 2009, we can in fact expect convergence of informant judgments and the use of informant judgments will no longer have to be regarded as an act of dishonesty.

(29) "If you make a theory, for example, and advertise it, or put it out, then you must also put down all the facts that disagree with it, as well as those that agree with it." (See the Feynman's remark quoted above.)

- Question: How do we know or how can we reasonably determine what qualifies as a "fact" in the context of research concerned with the properties of the Computational System?

4. Proposal

- My answer: *Repeatable phenomena*[15]

✓ **Claim**: A *repeatable phenomenon* is ***the most basic empirical unit of 'facts'*** in research concerned with the properties of the Computational System.

A *repeatable phenomenon* consists of a **Schema*-based prediction that has survived a rigorous test of disconfirmation and the corresponding ^{ok}*Schema*-based predictions that have been confirmed and hence it must consist of a **Schema* and the corresponding ^{ok}*Schemas*, and their corresponding **Examples* and ^{ok}*Examples*, and the informant judgments on those examples.

(30) A **Schema*-based prediction:
The informant judgment on α under interpretation γ(*a*, *b*) is "totally unacceptable" for *any* **Example* of a **Schema*.

(31) An ^{ok}*Schema*-based prediction:
The informant judgment on α under interpretation γ(*a*, *b*) is "fully acceptable" for *some* ^{ok}*Example* of an ^{ok}*Schema*.

(32) An ^{ok}*Schema*-based prediction, an alternative formulation:
The informant judgment on α under interpretation γ(*a*, *b*) is "not totally unacceptable" for *some* ^{ok}*Example* α of ^{ok}*Schema* σ.

What formulation of an ^{ok}*Schema*-based prediction to adopt depends in part on whether we are engaged in a *single-informant experiment* or a *multiple informant experiment* and it is closely related to *resourcefulness* issues.[16]

Among the main methodological proposals in Hoji 2009 are:

(33) a. There is a fundamental asymmetry between a **Schema*-based prediction and an ^{ok}*Schema*-based prediction in terms of the significance of their failure (to be borne out).
b. The informant intuition is more directly revealing about the properties of the Computational System if it is on the (un)acceptability of a sentence *under an interpretation involving two expressions* than if it is on "simple" (un)acceptability of sentences.

➢ (Apparent) counterexamples as alluded to in (23) (hence in (15)) vs. "counterexamples" to an alleged *repeatable phenomenon*[17]

[15] One of the chapter (chapter 2, which is called *Repeatable Phenomenon*) of Hoji 2009 is entirely devoted to an empirical illustration of the notion of *repeatable phenomena*.

[16] The issues are discussed in some depth in Hoji 2009: chapter 3.

[17] Lexical hypotheses about *zibunzisin* and *otagai* and alleged generalization regarding the "locality requirement" on a

--I suggest the heuristics in (34).[18]

(34) a. Build our theory about the Computational System and about the properties of items of the mental lexicon of the speakers of a language under discussion based on *repeatable phenomena*.
b. Do not use hypotheses about the Computational System or about properties of items of a mental lexicon if they have been shown not to be backed up by a *repeatable phenomenon*.

(35) "When you have put a lot of ideas together to make an elaborate theory, you want to make sure, when explaining what it fits, that those things it fits are not just the things that gave you the idea for the theory; but that the finished theory makes something else come out right, in addition." (See the Feynman's remark quoted above.)

--This corresponds to Lakatos' (1970/1978) *theoretically and empirically progressive problemshift*. If we fail to do the above, our work remains a description of 'facts'. It is the pursuit for *theoretically and empirically progressive problemshift* that has led to the insistence on having a tight connection between [a *repeatable phenomenon*] and [hypotheses about the Computational System and/or those about lexical items that the Computational System makes reference to]; cf. the model of prediction making in Hoji 2009 (and also in the handout for my 6/12 talk).

So, an alleged empirical generalization is evaluated not only in terms of (36a) but also in terms of (36b).

(36) a. whether it constitutes a *repeatable phenomenon*
b. whether the hypotheses that are claimed to be crucially related to, i.e., claimed to be responsible for, the empirical generalization, now "elevated to" the level of a *repeatable phenomenon*, contribute to making a new prediction concerning a *new repeatable phenomenon*

Question: But is it really feasible or even possible to pursue research in accordance with the methodological proposal alluded to above?
Answer: Hoji 2009 provides an illustration of how it can be pursued.

--But the research pursued in Hoji 2009 narrows down the empirical coverage considerably by adopting the Reinhartian heuristic in (22), repeated here.

(22) The Reinhartian heuristic:
The relation at LF that underlies γ(a, b) mentioned in a bridging statement must be based on a *c-command* relation between LF(a) and LF(b).

It is hoped that once we have obtained a firm understanding of a limited empirical domain by following the heuristic in (22), we can make crucial reference to the results we will have obtained and can expand the empirical coverage a great deal (if not exponentially), without losing the empirical testability of our hypotheses.

5. Some implications

The research concerned with properties of the Computational System is, or at least, can be, regarded as part of an attempt to understand what characterizes the human being, as opposed to other beings, organic and inorganic. As to the general significance of the proposed methodology, I would like to think/suggest that, if successful, the research being pursued here will show (37)-(39).

(37) The core properties of the language faculty can be investigated scientifically in line with (6), repeated here.

(6) The general scientific method:
Guess — Computing Consequence — Compare, Experiment

(38) a. The empirical merit of particular linguistic theorizing can, and hence, in my view, should, be determined by experiments.

---

numeral-classifier sequence come to mind, for example.

[18] There are also low-level issues concerning a minimal level of professionalism independent of the heuristics in (34), having to do with citation, for example, and such problems might also be related to the failure to adhere to heuristics such as (34).

b. The interpretation of the experimental results does not require statistics of much sophistication, in regard to the most crucial criterion in hypothesis evaluation.
c. *In principle*, anyone can be a judge on the validity of hypotheses about properties of the Computational System.

(39) Given (38), the field will (finally) be free from:
a. English-centricity (and other related "guidelines" that are seemingly accepted by many practitioners)
b. Authoritarianism

- English centricity
- Authoritarianism
- Degeneration into a metaphysical research program (despite the appearance of dealing with empirical issues and materials)

I would like to further speculate that *science* will become accessible to anyone, in principle, to the extent that anyone can participate in a scientific experiment and appreciate the significance of its results without a special talent or training in mathematics. Together with (39), this has the *potential* of having a non-trivial implication of fostering less reliance on authority in general. Under the proposed methodology, the only authority one would pay heed to is the result of an experiment; nothing else matters, ultimately, very much like what Richard Feynman advocated in his 1964 "Messenger Lectures" at Cornell University.

I should like to conclude the discussion on (39b) by considering some (political) remarks by Chomsky and Geoff Pullum's remark on Chomsky's research program; cf. also Kuroda 2008: section 8.[19]

(40) "Compare mathematics and the political sciences—it's quite striking. In mathematics, in physics, people are concerned with what you say, not with your certification. But in order to speak about social reality, you must have the proper credentials, particularly if you depart from the accepted framework of thinking. Generally speaking, it seems fair to say that the richer the intellectual substance of a field, the less there is a concern for credentials, and the greater is the concern for content."
(Chomsky 1979: 7 (*Language and Responsibility*))

(41) "The book is really excellent, I think: vitally needed, eminently readable, and right on the mark with its comprehensive and incisive critique of the most influential *confidence trick* in the history of modern linguistics..."
(Geoffery Pullum, on the back cover of *Chomsky's Minimalism* by Pieter A. M. Seuren, 2004, Oxford University Press) (Emphasis by HH.)

(42) "Now, if you ask, "What media can I turn to to get the right answers?" First of all, I wouldn't tell you that because I don't think there's an answer. The right answers are what *you* decide are the right answers. Maybe everything I'm telling you is wrong. Okay? Could perfectly well be; I am not God. But that's something for *you* to figure out. I mean I can tell you what *I* think happens to be more or less right. But there isn't any reason why you should pay any attention to it."
(Noam Chomsky, in *Manufacturing Consent: Noam Chomsky and the Media* (1992)[20])

湯川秀樹　理論物理学の輪郭　in「目に見えないもの」 (first published in 1946, the page references are to the 1976 講談社学術文庫.)

経験事実との比較によって、この両説の当否を決定することは不可能であったから、原子論と連続論とは長い間対立して存続し得たのみならず、原子論の側においては原子自身が人間の肉眼には見えない想像の産物であったから——各民族の神話の間に異同があるごとく——たがいに類似し、あるいはたがいに相違する種しゅの原子模型が考え得たのである。これらの多くの学説の栄枯盛衰を左右したのは、主唱者ないし追随者の人格や見識に対する信頼の程度とか、表現の巧拙とか、宗教的権威とかいう多かれ少なかれ個人的ないし社会的な幾つかの因子であった。(pp. 12-13)

[19] Kuroda, S.-Y. 2008. "Mathematics and Generative Grammar—"Beyond Explanatory Adequacy" and Mathematical Realism of Language: A Fable for Naoki Fukui," *Sophia Linguistica 56*. pp. 1-36. Those interested in obtaining a copy of the paper should contact Prof. Yukio Otsu.

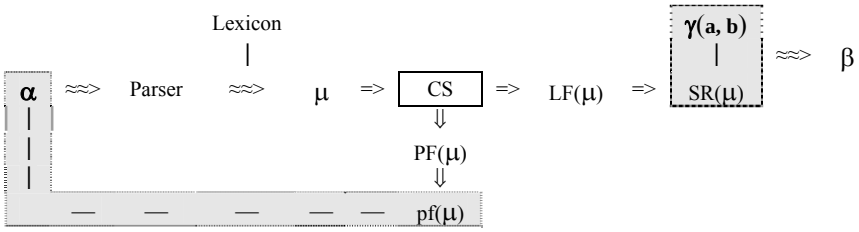
[20] The video can be downloaded at:
http://video.google.com/videoplay?docid=-5631882395226827730
The subject matter is not linguistics but it is highly recommended.

Suppose that someday the methodology advocated in Hoji 2009 becomes the norm of the field. Anyone can place on-line his/her hypotheses (necessarily with some relevant generalizations), along with the prediction(s) (both **Schema*-based predictions and ^{ok}*Schema*-based predictions), the experimental specifications (including preliminary experiments) (including the **Schemas* and the corresponding ^{ok}*Schemas*) and perhaps actual experiments (hence including actual **Examples* and ^{ok}*Examples*) and their results. And they can have a reasonable expectation to be taken seriously irrespective of his/her credentials as long as their **Schema*-based prediction(s) has/have survived a rigorous test of disconfirmation and the corresponding ^{ok}*Schema*-based predictions have been confirmed.

Perhaps, it is not unreasonable to think that we are pursuing the methodology along the lines with Hoji 2009, which is Popperian in its essentials, because we do not want to be judged by our credentials. That is to say, we do not want to be part of the "game" where our work is judged by whether we have the blessing of the authority (e.g., cited by so and so, published in such and such journals or from such and such publishers, etc.), whether the majority of the subfield you belong to agrees with you, etc.[21]

6. Appendix I:

(43) The Model of Judgment Making by the Informant on **the acceptability of sentence α with interpretation γ(a, b)** (due to A. Ueyama):



a. α: presented sentence
b. γ(*a*, *b*): the interpretation intended to be included in the 'meaning' of α involving expressions *a* and *b*[22]
c. μ: numeration
d. LF(μ): the LF representation that obtains on the basis of μ
e. SR(μ): the information that obtains on the basis of LF(μ)
f. PF(μ): the PF representation that obtains on the basis of μ
g. pf(μ): the surface phonetic string that obtains on the basis of PF(μ)
h. β: the informant judgment on the acceptability of α under γ(*a*, *b*)

(44) The model of quantifying the informant judgment β on the acceptability of sentence α under interpretation γ(a, b) (due to Y. Deguchi):

β ranges between 0 and 1, with the former corresponding to 'complete unacceptability' while the latter corresponding to 'full acceptability'.

β = [G] − [P] − [I], where

[G] is 1 if and only if (i) PF(μ) obtains[23] and (ii) SR(μ) compatible with γ(*a*, *b*) obtains; otherwise, [G] is 0.

[P] (0 ≤ [P]) represents the degree of difficulty the informant 'feels' in 'obtaining' or trying to 'obtain' μ, as it is

---

[21] If we accepted that it is okay for us to proceed as in Chomsky's "As for my own methods of investigation, I do not really have any. The only method of investigation is to look hard at a serious problem and try to get some ideas as to what might be the explanation for it, meanwhile keeping an open mind about all sorts of other possibilities. Well, that is not a method. It is just being reasonable, and so far as I know, that is the only way to deal with any problem, whether it is a problem in your work as a quantum physicist or whatever" (see (11) above), we will most likely end up having to play the game alluded to above, and that seems to be exactly what has happened in generative grammar up to now in my opinion To be fair to Chomsky, we should consider the above remark of his as referring to *the context of discovery*. If so, we agree with him. But then, the problem is that he does not seem to have, or at least has not proposed, articulated, or developed, any method for *the context of testing* (often called *the context of justification* but one might call it *the context of falsification*).

[22] See (8b) and (9b) in the "Examples handout," for instance.

[23] It is assumed that the informant's *string sensitivity* ensures that PF(μ) and α are identical. The *string sensitivity* of the informant has to be ensured by preliminary experiments, for example.

reflected in β.

[I] ($0 \leq$ [I]) represents the degree of unnaturalness the informant 'feels' about SR(μ) compatible with γ(*a*, *b*), as it is reflected in β.

In terms of 'detectable effects', what *would* be β≤0 is the same as β=0; i.e., β=0 and what *would* be β≤0 are both regarded as expressing 'total unacceptability', represented here as β=0.

(44) expresses the observation that the difficulty in 'parsing' the presented sentence α "as intended" (i.e., so as to yield the intended LF representation and the unnaturalness of the entire SR(μ)) may result in the informant judgment that α is not fully acceptable (or even totally unacceptable) under γ(a, b)).

(45) (Cf. (30).)
A **Schema*-based prediction[24]:
The informant judgment on α under interpretation γ(*a*, *b*) is *always* β=0 for *any* **Example* of a **Schema*.

(46) (Cf. (31).)
An ^{ok}*Schema*-based prediction:
The informant judgment on α under interpretation γ(*a*, *b*) is β=1 for *some* ^{ok}*Example* of an ^{ok}*Schema*.

(47) (Cf. (32).)
An ^{ok}*Schema*-based prediction, an alternative formulation:[25]
The informant judgment on α under interpretation γ(*a*, *b*) is 0<β for *some* ^{ok}*Example* of ^{ok}*Schema*.

- the informant's *resourcefulness*
- *single-informant experiments* vs. *multiple-informant experiments*

(48) Confirmability and disconfirmability

	Confirmation	Disconfirmation
^{ok}*Schema*-based predictions	possible	impossible
**Schema*-based predictions	impossible	possible

REVIEW:
A **Schema*-based prediction is β=0 because it is predicted that [G]=0 for any **Example* α of any **Schema* and that should result in β=0. The judgment that α is not totally unacceptable under γ(*a*, *b*) (even if not fully acceptable) would therefore disconfirm a **Schema*-based prediction. Notice that such a judgment should mean that, corresponding to α, there *is* SR(μ) compatible with γ(*a*, *b*). This in turn should mean [G]=1 and the actual value of β being lower than 1 must be due to [P] and/or [I]. While the *marginal acceptability* would disconfirm a **Schema*-based prediction, as just noted, it would be compatible with an ^{ok}*Schema*-based prediction in (47) since, according to (44), 0<β≤1 must mean [G]=1.

We suggest that the informant judgments must obtain as indicated in (49) in order for a *repeatable phenomenon* to obtain involving the bridging statement that has given rise to the **Examples* and the corresponding *okExamples*.

(49)

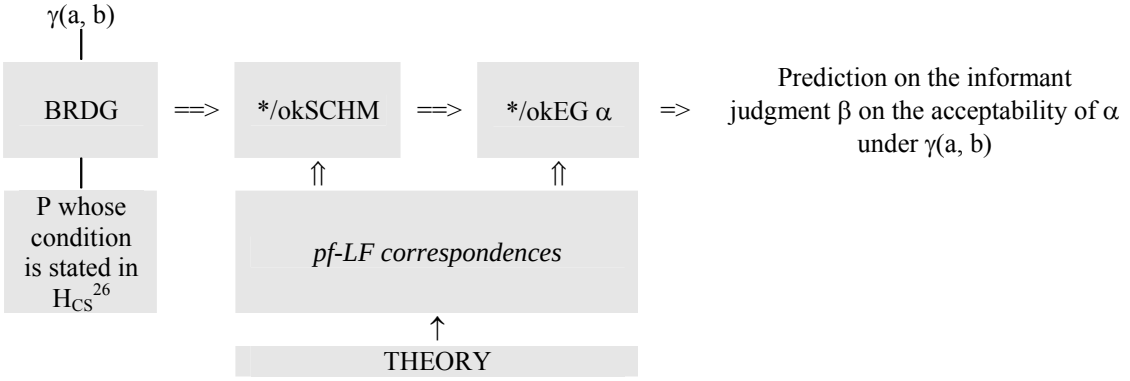
	the judgments *necessary* for a *repeatable phenomenon* to obtain
**Examples*	β=0
Corresponding ^{ok}*Examples*	0<β≤1

- What is *most crucial* for a *repeatable phenomenon* to obtain is the 'values' of the **Schema* and those of its corresponding ^{ok}*Schemas*, which would obtain on the basis of the informant judgments on the **Schema* and the ^{ok}*Schemas*.

[24] See (5) and (6) in the "Examples handout," for instance.

[25] Other formulations of an ^{ok}*Schema*-based prediction are also considered in Hoji 2009.

(50) The model of prediction making:



a. BRDG: the *bridging statement* that relates γ(a, b) and P, by specifying P as a necessary condition, stated in theoretical terms, for γ(a, b).[27]
b. P: the property at LF mentioned in BRDG[28]
c. H_{CS}: the hypothesis about the Computational System that states the condition for P. (Universal)
d. */okSCHM: *Schema or okSchema
e. */okEG: *Example (in the case of ****Schema***) or okExample (in the case of ^{ok}***Schema***)
f. H_{Lex}: the hypothesis about an item in the mental Lexicon.[29] (Language specific)
g. ***pf-LF correspondences***: (the general patterns of) the pf-LF correspondences (assumed by the researcher)
h. THEORY: the hypotheses adopted elsewhere in the theory (i.e., other than the H_{CS} under discussion), including those about the Computational System and those about items of the mental Lexicon of the speakers of the language in question.

What does each of (51) mean/indicate?

(51) a. a ****Schema***-based prediction's getting disconfirmed
b. a ****Schema***-based prediction's surviving a rigorous attempt of disconfirmation
c. an ^{ok}***Schema***-based prediction's getting confirmed
d. an ^{ok}***Schema***-based prediction's failing to be confirmed

- We cannot meaningfully address the significance of each of (51) in isolation, except for (51a).
- When a ****Schema***-based prediction on sentence α under γ(*a*, *b*) has survived a rigorous test of disconfirmation, how could we ensure that the total unacceptability of the ****Examples*** is indeed due to the properties of the H_{CS} (and/or H_{LEX}), not due to some parsing difficulty of some magnitude?
- By making sure that the same surface forms as those ****Examples*** are not totally unacceptable if we eliminate the consideration about γ(*a*, *b*).
- If a ****Schema***-based prediction were on simple unacceptability of sentence α, we would not have a similar means to ensure that the total unacceptability of the ****Examples*** is indeed due to the properties of the H_{CS} (and/or H_{LEX}) under discussion.
- Hence (33b), repeated here.

(33) b. The informant intuition is more directly revealing about the properties of the Computational System if it is on the (un)acceptability of a sentence *under an interpretation involving two expressions* than if it is on "simple" (un)acceptability of sentences.

---

[26] It is necessary to add here "and/or H_{LEX}"; see (50f). The model in (50), however, does not make reference to H_{LEX} because H_{LEX} can affect ***pf-LF correspondences*** and its inclusion would complicate the presentation of the model here.

[27] It is argued in Hoji 2009: chapter 3, section 4.1.1 that the *bridging statement* must be of the form in (i) as long as we are dealing with the informant judgment on the acceptability of sentences under interpretation γ(*a*, *b*).

(i) A certain linguistic intuition such as the sense of the availability of interpretation γ(*a*, *b*) arises only if (a) certain condition(s) is/are met *at LF*.

[28] *FD*, discussed in Hoji 2009, is an instance of *P* here.

[29] This is not included in the above chart.

7. Appendix II: Recommended audio-visual materials and readings

(52) YouTube videos:
   a. "Feynman on Social Sciences," (1' 52") http://www.youtube.com/watch?gl=JP&hl=ja&v=_EZcpTTjjXY
   b. "Feynman"
      Part 1: http://www.youtube.com/watch?v=knDXAr4ltMA (9' 06")
      Part 2: http://www.youtube.com/watch?v=3U0PPunfYFA (9' 15")
      Part 3: http://www.youtube.com/watch?v=qILEIio8jZU (9' 12")
      Part 4: http://www.youtube.com/watch?v=wd2T50tWOfY (9' 07")
      Part 5: http://www.youtube.com/watch?v=Fsz9CLmNTJI (9' 07")
      Part 6: http://www.youtube.com/watch?v=eJ3CkWzpyTo (9' 07")
   c. YouTube videos of Feynman lectures you can find under "Messenger Lectures."

(53) Audio:
      Lakatos "Science and pseudo-science"[30]

(54) Papers:
   a. Popper, Karl. 1963. "Science: Problems, Aims, Responsibilities," Federation Proceedings (Baltimore), Federations of American Societies of Experimental Biology Vol. 22, Issue 4: 961-972.
   b. Richard Feynman. 1974 "Cargo Cult Science," a Caltech Commencement Address.[31] (Reproduced (with slight adaptation) in *Surely You're Joking, Mr. Feynman!*)

[30] The audio tape and its written version—which appears in Lakatos 1978 as "Introduction: Science and Pseudoscience." (pp. 1-7)—are available at: http://www.lse.ac.uk/collections/lakatos//Default.htm.

[31] The paper is available on-line. Two of the URLs of the paper are:
http://www.columbia.edu/itc/applied/wiggins/Classes/E4903/Fall2003/cargo.pdf
http://www.physics.brocku.ca/etc/cargo_cult_science.php