June 12, 2009
Tohoku University

Toward A Scientific Study of the Language Faculty: a proposal and implications

Hajime Hoji
University of Southern California
http://www.gges.org/hoji/
hoji@usc.edu

- 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 talk addresses: (i) the general method in science (that I find reasonable), (ii) the background and the proposal, and (iii) the significance and implications of the proposal, not necessarily in this order.
- Because of the time constraints, I regret that the talk may not provide (almost) any concrete illustration of the proposal.[1]

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.*[2]

(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?

---

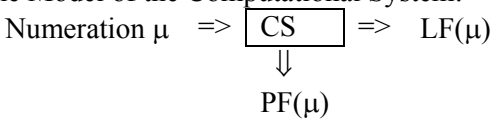
[1] If you would like to learn (more) about the specific aspects of the proposal in Hoji 2009 and their empirical illustration, including how specific hypotheses get revised or rejected in accordance with its methodological proposal, please email me at *hoji@usc.edu*.

[2] There seems to be a typo here. But I am reproducing what is in the volume because what is intended seems clear enough.

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'.

(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

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. 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." The main goal of *generative grammar* can thus be understood as demonstrating the existence of such an algorithm by discovering its properties. Construed in this way, it is not language as an external 'object' but the *language faculty* that constitutes the object of inquiry in generative grammar.

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

(4) 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 (4a))

CONCLUSION AT THIS POINT:

- The object of inquiry in generative grammar is what underlies the language users' intuitions about the relation between "sounds" (and signs in a sign language) and "meanings."

  - What does that mean for the use of the researcher's own introspective judgments in hypothesis formation and hypothesis testing? (This question raises a number of interesting methodological issues, concerning *repeatability*, cross-linguistic research, etc., and how one answers the question may be a good indication about one's research 'orientation'.)

2. The Main idea of this talk

I would like to 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.

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

YouTube video: http://www.youtube.com/watch?v=knDXAr4ltMA
Feynman Part 1 (about 3' 15" from the beginning of the tape) (See 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, I will propose that the language faculty can be studied with the general scientific method as schematized in (5) and make concrete suggestions as to how.

In order to proceed with our investigation of the language faculty with the general scientific method in (5), we must ensure the following, at least to a minimally satisfactory degree.

(6) 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.

(6a) and (6b) can in turn be ensured only if (7) and (8) hold, respectively.

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

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

Furthermore, we should ensure (9) as best as we can.

(9) 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] 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.)

3. Proposal

3.1. The main proposal

The main proposal to be discussed in this talk is (10).

(10) If we want to discover the properties of the Computational System that is hypothesized to be at the center of the language faculty, what I call a *repeatable phenomenon* should be considered as the *minimal unit of 'facts'* for such research; see also (41) below.

A PREVIEW

What is meant by a *repeatable phenomenon*:

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.

(11) A **Schema*-based prediction:
The informant judgment on α under interpretation γ(*a*, *b*) is *always* β=0 (i.e., totally unacceptable) for *any* **Example* of a **Schema*.

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

(13) An ^{ok}*Schema*-based prediction, an alternative formulation:
The informant judgment on α under interpretation γ(*a*, *b*) is 0<β (i.e., not totally unacceptable) for *some* ^{ok}*Example* α of ^{ok}*Schema* σ.

In regard to (8), the consequences of the "guess" are to be related, ultimately, to a set of observations that would or would not constitute a *repeatable phenomenon*.

3.2. Some specific aspects of the proposal

I would like to try to go over the following aspects of the proposal.

(14) a. There is an 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.

3.2.1. The model of judgment making

➢ For (14a), let us turn to **the model of judgment making** in (15).

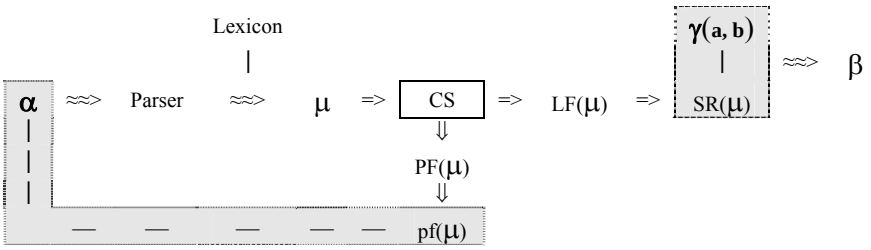
The essential aspects of (15) are an immediate consequence of adopting the thesis that the Computational System (see (3)) is at the center of the language faculty.

A hypothetical dialogue between a typical "generative grammarian" (G) and a naïve outsider (N)[4]:
N: So, you are interested in what underlies the language users' intuitions about the relation between "sounds" (and signs in a sign language) and "meanings."
G: Yes.

[4] Hoji 2009: Appendix compares (15) with the model of judgment making suggested in Schütze 1996. Schütze, Carson. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*, University of Chicago Press.

N: I suppose you construct, or at least test, your hypotheses on the basis of the language users' intuitions.
G: Yes.
N: You should then have some idea about what goes on when the language users or your informants judge a sentence.
G: Well, ...
N: What is then your hypothesis about what the informant does when s/he judges a sentence?
G: ...

(15) The Model of Judgment Making by the Informant on **the acceptability of sentence α with interpretation γ(a, b)**[5] (due to A. Ueyama):



a. α: presented sentence
b. μ: numeration
c. γ(a, b): the interpretation intended to be included in the 'meaning' of α involving expressions *a* and *b*[6]
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)

(16) 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[7] and (ii) SR(μ) compatible with γ(a, b) obtains; otherwise, [G] is 0.
[P] ($0 \leq [P] \leq 1$) represents the degree of difficulty the informant 'feels' in 'obtaining' or trying to 'obtain' μ, as it is reflected in β.
[I] ($0 \leq [I] \leq 1$) represents the degree of unnaturalness the informant 'feels' about SR(μ)

---

[5] A numeration is an input to the CS and its output representations are LF and PF, and that is indicated by "==>" in (15). The two arrows before and after *CS* in (15) thus represent the 'is the input of' and the 'yields as an output' relations, respectively. Similarly, what is meant by the arrow between *LF* and *SR* is that *SR* obtains based on *LF*. What is intended by" ≈≈>," on the other hand, is not an input/output relation and are used more loosely, as indicated in (i).
(i) a. Presented Sentence ≈≈> Parser: ... is part of the input to ...
b. Parser ≈≈> Numeration: ... contributes to the formation of ...
c. SR ≈≈> Judgment: ... serves as *a* basis for ...

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

[7] 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.

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.

The merit of understanding the informant judgment as β = [G] − [P] − [I]:
The informant cannot tell us the value of [P] or [I]. Despite the indeterminacy of [P] and [I], we know that β=0 if [G]=0, given that β = [G] − [P] − [I]. We capitalize on that. We also recognize that β can vary a great deal between 0 and 1 if [G]=1. That is also part of our consideration in characterizing what we should consider as a *repeatable phenomenon*; the methodology put forth in Hoji 2009 is motivated not only by (17a) but also (17b).

(17) a. the conception we adopt about the nature of the language faculty and the Computational System embedded in it and how hypotheses about the Computational System can be tested
b. our concern about articulating and specifying the nature of the fallibility of our hypotheses as clearly as possible so as to maximize our chances of learning something about the Computational System from "failed predictions"

(16) 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)).

(18) (Cf. (11).)
A **Schema*-based prediction[8]:
The informant judgment on α under interpretation γ(*a*, *b*) is *always* β=0 for *any *Example* of a **Schema*.

(19) (Cf. (12).)
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*.

(20) (Cf. (13).)
An *ok*Schema*-based prediction, an alternative formulation:[9]
The informant judgment on α under interpretation γ(*a*, *b*) is 0<β for *some ok*Example* α of *ok*Schema* σ.

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

(21) 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)

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

[9] Other formulations of an *ok*Schema*-based prediction are also considered in Hoji 2009.

(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 (20) since, according to (16), 0<β<1 must mean [G]=1.

3.2.2. The claim

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

We suggest that the informant judgments must obtain as indicated in (22) in order for a *repeatable phenomenon* to obtain involving the bridging statement that has given rise to the **Examples* and the corresponding *okExamples*.

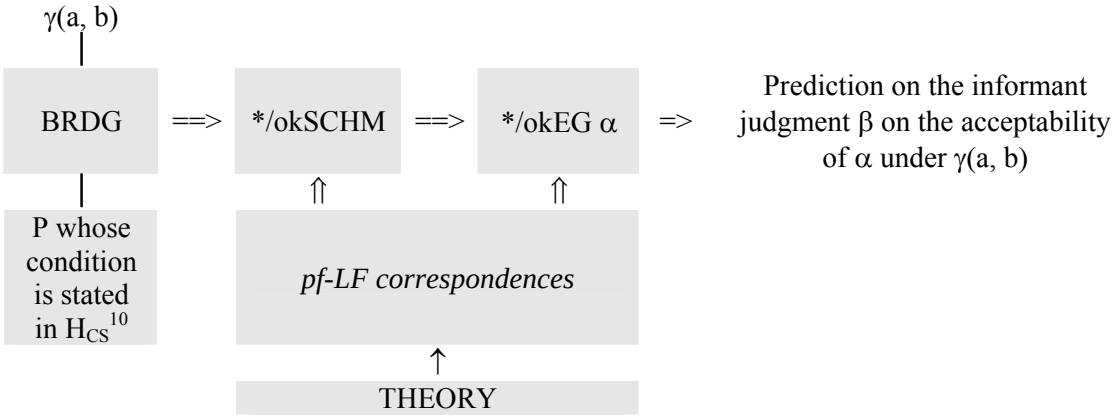
(22)

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

3.2.3. The model of prediction making

➢ For (14b), let us turn to **the model of prediction making** in (23).

(23) 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).[11]

---

[10] It is necessary to add here "and/or H_{LEX}"; see (23f). The model in (23), 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.

[11] 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*.

It is also noted there that the *bridging statement* is not testable unless the hierarchical relation 'gets converted' to a precedence relation because the informant cannot detect hierarchical relations among elements in the presented sentence. We thus need to 'convert' the *bridging statement* to a statement of the form in (ii).

b. P: the property at LF mentioned in BRDG[12]
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.[13] (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 (24) mean/indicate?

(24) 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 (24) in isolation, except for (24a).
- 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 (14b), repeated here.

(14) 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.

- I thus suggest (25).

(25) The **Minimum Paradigm Requirement**: *when working under* (15), a paradigm must minimally consist of examples of the following three types
a. a *Example such that at least one of the conditions (structural or lexical) for γ(a, b) is *not* satisfied in any of the LF representations that could correspond to it
b. an okExample$_1$ such that it minimally differs from (25a) and the structural and lexical condition(s) for γ(a, b) is/are satisfied in an LF representation that could correspond to it
c. an okExample$_2$ such that it is identical to (25a) in terms of the surface string but with an interpretation that does not include γ(a, b)

(ii) A certain linguistic intuition such as the sense of the availability of interpretation γ(*a*, *b*) arises only if *a* and *b* appear in a *linearly arranged schema* of a particular form.

We can make a statement of the form in (ii) empirically testable only if we commit ourselves to particular ***pf-LF correspondences***.

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

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

3.3. Learning from errors

What crucially underlies the contention (made in Hoji 2009) that "an alleged generalization that does not form a *repeatable phenomenon* has not (yet) attained the status of *data in generative grammar*" is the desire to accumulate hypotheses about the Computational System that are empirically testable and to proceed in accordance with the general heuristic in (26).

(26) (Chapter 5: (19))
A general research heuristic:
We should maximize our chances of learning something about the properties of the Computational System from the disconfirmation of our predictions.

(26) is very much along the lines of what Popper (1963) provides as his own summary of the theses that he has put forth. I only copy the first 4 of the 17 points that Popper (1963: 965-966) gives, "re-stating all the controversial things I have been saying in a number of theses which I shall try to put in as challenging a form as I can."[14]

(27) (Popper's (1963: 965) (1)-(4))
a. All scientific knowledge is hypothetical or conjectural.
b. The growth of knowledge, and especially of scientific knowledge, consists in learning from our mistakes.
c. What may be called the method of science consists in learning from our mistakes systematically; first, by daring to make mistakes—that is, by boldly proposing new theories; and second, by searching systematically for the mistakes we have made, that is, by the critical discussion and the critical examination of our theories.
d. Among the most important arguments which are used in this critical discussion are arguments from experimental tests.

The insistence on emphasizing the significance of **Schema*-based prediction and the postulation of the model of prediction making in (23) have been prompted by the consideration in (26); see (14a). So is the insistence on relying on the informant judgments on the *(un)acceptability of sentence* α *under interpretation* γ*(a, b)*, see (14b).[15]

3.4. Some empirical illustration

See the separate handout: the handout *would* include some Japanese sentences that have been used in our preliminary experiments, and they, along with the "judgments" by the present audience, may serve as a basis of illustration of (14) (and perhaps other notions such as *single-informant experiment*, *multiple-informant experiment*, *structural/contextual resourcefulness*, etc.) [I have the

---

[14] See Popper 1963: 965-966. (Popper, Karl. 1963. "Science: Problems, Aims, Responsibilities," *Federation Proceedings (Baltimore), Federations of American Societies of Experimental Biology* Vol. 22, Issue 4: 961-972.) Feynman's basic points/theses (as in his 1974 Caltech address and in his 1964 Cornell Lectures) seem strikingly similar to Popper's. A similar point is already made in Poincaré 1902 *LA SCICENCE ET L'HYPOTHESE*. The page reference below is to its 1952 English translation *Science and Hypotheses* (Dover Publications).
"… The physicist who has just given up one of his hypotheses should, on the contrary, rejoice, for he found an unexpected opportunity of discovery. His hypothesis, I imagine, had not been lightly adopted. It took into account all the known factors which seem capable of intervention in the phenomenon. If it is not verified, it is because there is something unexpected and extraordinary about it, because we are on the point of finding something unknown and new. Has the hypothesis thus rejected been sterile? Far from it. It may be even said that it has rendered more service than a true hypothesis. Not only has it been the occasion of a decisive experiment, but if this experiment had been made by chance, without the hypothesis, no conclusion could have been seen; and only one fact the more would have been catalogued, without deducing from it the remotest consequence."(Poincaré 1952: chap. 9, 150-151)

[15] Hoji 2009: chapter 5, section 2 contains further discussion.

BVA, including the OS, paradigms in mind, not the *zibunzisin/otagai* paradigms because this talk is not concerned with bad research practice in the field. More in particular, I have in mind paradigms on: *a* vs. *so*, the 'split antecedence', the basic BVA paradigm, including OSV, and "resumption." The Principle-B-related materials, which are in a way the empirical climax in Hoji 2009, are not included in the separate handout.]

3.5. Summary

The proposal in Hoji 2009 can be understood as being driven by the desire to discover the properties of the language faculty with the scientific method schematized in (5), along with the research heuristic in (26). As indicated above, (5) and (26) are not novel ideas by any means in regard to scientific practice. The question is, really, whether we can (aspire to) discover the properties of the language faculty by 'following' (5) and (26). I suggest in Hoji 2009 that we can, but only if we start with something that we can manage at our initial stage of investigation; see (28).

(28) 科学というものには、本来限界があって、広い意味での再現可能の現象を、自然界から抜き出して、それを統計的に究明していく、そういう性質の学問なのである。（中谷宇吉郎「科学の方法」, 1958, 17, as cited in Uchii 1995: 56「科学哲学入門」[16])

(10) If we want to discover the properties of the Computational System that is hypothesized to be at the center of the language faculty, what I call a *repeatable phenomenon* should be considered as the *minimal empirical unit of 'facts'* for such research.

4. General significance and 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 (29)-(33).

(29) The core properties of the language faculty can be investigated scientifically in line with (5), repeated here.

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

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

---

[16] Uchii's (1995) point is not (necessarily) to endorse the view in (i); he makes reference to it in the last of his chapter 2, which he concludes by stating that methods in science can change *rekisitekini*, and in accordance with the change, the characterization of (ii) can also change.

(ii) a. what is scientific knowledge
b. what is science

To the extent that それを統計的に究明していく is part of what counts as science, which incidentally seems to be a fairly common view among certain researchers—it may indeed be part of the 'definition of science today' and some people seem to think that we must mimic it in linguistics as well. Hoji 2009, on the other hand, advocates the possibility that a scientific study of the Computational System can be pursued without necessarily requiring statistics of much sophistication although we do need "basic statistics" (such as the averaging of "scores" among the informants) to measure the reproducibility of the result of an experiment.

While it is not entirely clear how *repeatability* could be measured in the context of cross-linguistic empirical research, it seems useful to consider the issue in light of the thesis in (31).

(31) *Across-speaker repeatability* can be meaningfully addressed only if *within-speaker repeatability* (*across-occasion* and *across-example repeatability*) obtains.

(32) A cross-linguistic empirical claim can be meaningfully addressed only if *within-speaker repeatability* obtains in regard to the issue/phenomenon in the language(s) under discussion, and it would in fact be well to have also achieved some degree of *across-speaker repeatability*.

In other words, it seems rather senseless to address a cross-linguistic empirical claim without having obtained *within-speaker repeatability* and *across-speaker repeatability* in each of the languages under discussion. This seems rather common-sensical. But the point is perhaps worth making in light of the fact that a cross-linguistic study seems to often make crucial reference to an alleged generalization that falls *far short* of being a *repeatable phenomenon*, as in the case of the alleged generalizations discussed in Hoji 2009 and elsewhere regarding *zibunzisin* and *otagai* in Japanese.

Convincing the others (presumably the other practitioners in the field if not those outside the field), I believe, is part of science. Obtaining *repeatability* is a necessary condition for convincing the others. And that makes it imperative that we develop a reliable experimental methodology to test the validity of one's hypotheses and especially a reliable method of evaluating the result of an experiment. But such methods have a function beyond convincing the others. It also has the function of making us feel willing to be convinced by others.

One might find the point rather obscure if one only thinks about interaction among the native speakers of one's own language(s). Suppose one is evaluating someone else's work that deals with a language that one does not speak as one's native language. One can never be sure about the reliability of the generalizations presented in such work. What does one do then? Well, some people may simply assume that the presented generalizations are valid, i.e., that they assume that they are *repeatable phenomena* in the terms of the preceding discussion. Some people may do so only if the alleged generalizations would support what they are pursuing; and this seems to be a rather typical practice in the field as far as I can tell. Others may think like the following: "Well, maybe valid; but maybe not. So I will take them as valid only if I detect something analogous in my own language, and until then I leave them in the category of 'Maybe'." Now, we would have a rather different attitude if the alleged generalization were presented along with the relevant experiment(s) and its/their result(s)—which would presumably include the **Schema(s)*, the *^{ok}Schemas*, and **Examples*, and the *^{ok}Examples*. We would in that case be much more willing to accept the proposed generalizations as valid, insofar as they form a *repeatable phenomenon*.

Accepting such a *repeatable phenomenon* as being established in regard to another language might in fact help us with our research on our own language since we would in that case have good reason to believe that, unless there is reason otherwise, the same generalization should hold in our own language, provided that the generalization is based on a universal statement. Much of cross-linguistic research, however, seems to proceed without being seriously concerned with whether an alleged generalization constitutes a *repeatable phenomenon* and we sometimes, if not often, see an alleged generalization adopted despite a demonstration in published works that it clearly fails to qualify as a *repeatable phenomenon*.

(33) Given (30), the field will (finally) be free from:
- a. English-centricity (and other related "guidelines" that are seemingly accepted by many practitioners)
- b. Authoritarianism

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 (33), 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; see the Feynman remarks quoted at the beginning of the handout (right after (5)).

I should like to conclude the discussion on (33b) by considering some (political) remarks by Chomsky and Geoff Pullum's remark on Chomsky's research program; cf. also Kuroda 2008: section 8.[17]

(34) "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*))

(35) "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..." (Emphasis by HH.)
(Geoffery Pullum, on the back cover of *Chomsky's Minimalism* by Pieter A. M. Seuren, 2004, Oxford University Press)

(36) "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)[18])

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)

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

---

[17] 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. Masa Koizumi.

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

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 of 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.[19]

5. A Summary in the form of questions

(37) In regard to research concerned with the properties of the language faculty:
- a. What are our "guesses," i.e., hypotheses, about?
- b. How could we obtain (i.e., *compute*) *consequences* from the hypotheses in question?
- c. How can such consequences be tested and what criteria could we have for determining whether the computed *consequences* agree or disagree with experiments?
- d. If our hypotheses are about the properties of the language faculty, what should such hypotheses look like, what primitive concepts and relations are likely to be included in the theory in terms of which such hypotheses are stated?
- e. How can they be tested in an experiment, what kind of experiments are possible and/or suitable to test the validity of such hypotheses?
- f. How are we to interpret the results of our experiments?
- g. When is our prediction disconfirmed or when are we willing to accept that it has been disconfirmed?
- h. When is our prediction confirmed or when are we justified to say that it has been confirmed?
- i What are the implications of disconfirmation of a prediction?
- j. How could we proceed in the event our prediction fails to be confirmed?

Questions like the following also come to mind.

(38) Can the data in generative grammar be categorical, i.e., clear-cut, to begin with?
Are the data in generative grammar necessarily probabilistic and hence require a statistic analysis in a crucial way?
Why do we need to conduct an experiment and why could we not test our hypotheses by

---

[19] If we accepted that it is okay for us to proceed as in Chomsky's (1988 *Language and Problems of Knowledge: The Managua Lectures*, MIT Press: 190) "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," we will most likely end up having to play the game just alluded to, 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*). (More on this in my 6/15 talk "Generative Grammar as an Empirical Science: its goal and how it can be pursued—with some historical reflections.")

observing how people use the language?[20]

6. Appendix I: Restated in the terms of Lakatos' *scientific research programs*

[from Richard Feynman's 1974 Caltech Commencement Address]

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.

(39) "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.)

--But how do we know what are "facts" in the context of research concerned with the properties of the Computational System?
--My answer: *repeatable phenomena*

---

[20] This is related to a much more general question of how a study of the Computational System is related to other types of studies of language and languages. Our answers depend upon what kind of hypotheses we would be considering, how we are to test their empirical consequences, and what we would take to be a disconfirmation of our prediction.

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

(40) "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 of 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 (23).

So, an alleged empirical generalization is evaluated not only in terms of (41a) but also in terms of (41b).

(41) 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*

7. Appendix II: Some remaining issues (and remarks of clarification)

(42) Some 'historical' clarification':
Q: So, did you develop your ideas based on Popper, Feynman, Lakatos, etc.?
A: Actually, no. I learned about what they say after having done my syntactic research for some time and formulated a(n) (vague) idea about how I think research concerned with the discovery of the properties of the language faculty ought to proceed. When I read what they say, I realized that there were people who had been concerned with many of the issues that I had been trying to address.

(43) Some clarification about the role of *repeatable phenomenon* in a scientific research program and about (what I take to be) the current stage of our research program:
a. The attempt to establish *repeatable phenomena* (in part by specifying how to do so) and to articulate its significance may be considered as what *has to* be done if *generative grammar* is ever to become a research program that deserves to be called an empirical science.
b. *Repeatable phenomena* are like "basic facts."
c. It is perhaps safe to assume that we are still at a pre-scientific stage of our research program, assuming that at its scientific stage we would be concerned more with how to deduce from our theory the number of *repeatable phenomena* (which would likely have attained a highly abstract character although they would ultimately be related to *repeatable phenomena* of the sort that we are trying to establish right now) than with establishing a *repeatable phenomenon*.
d. It is up to the researchers concerned with the discovery of the properties of the language faculty whether the field can eventually attain a status of an empirical science or remains to be a metaphysical research program, in which seemingly empirical materials and issues are addressed but it is never explicitly stated how the consequences of the predictions can be tested.

(44) A lingering/remaining question[21]:
Question: Is physics a right place to turn to as a "model" for a research program that is concerned with the language faculty?

8. Appendix III: Recommended audio-visual materials and readings

(45) 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."

(46) Audio:
Lakatos "Science and pseudo-science"[22]

(47) Papers:
Popper, Karl. 1963. "Science: Problems, Aims, Responsibilities," Federation Proceedings (Baltimore), Federations of American Societies of Experimental Biology Vol. 22, Issue 4: 961-972.
Richard Feynman. 1974 "Cargo Cult Science," a Caltech Commencement Address.[23] (Reproduced (with slight adaptation) in *Surely You're Joking, Mr. Feynman!*)

[21] I had been concerned with this question and was reminded of it by Y. Deguchi (p.c., May 2009).

[22] 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.

[23] 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