Model of Judgment Making and Hypotheses in Generative Grammar

Ayumi Ueyama[1]
Kyushu University, JAPAN

1. Introduction: the need of methodology

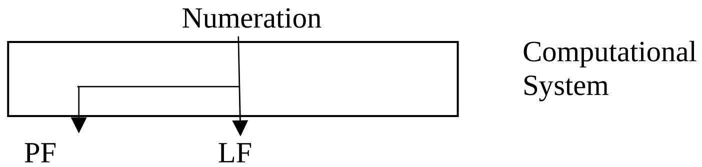
It is sometimes claimed that generative grammar does not need specific methodology because, being a branch of natural science, it should be governed by general principles of reasoning and nothing more is necessary. In this paper, we suggest, contrary to this long-standing view, that steady progress cannot be expected without some heuristic conventions for identifying what counts as evidence in the evaluation of hypotheses in generative grammar. We put forth a concrete means to do that, the adoption of which we maintain is a consequence of (i) accepting the model of the Computational System adopted here, (ii) committing ourselves to making our hypotheses about the Computational System empirically testable on the basis of the informants' acceptability judgments, and (iii) having the desire to ensure progress toward the goal of discovering the properties of the Computational System by making our hypotheses empirically testable.

2. Grammaticality and acceptability

2.1. Competence vs. performance

Generative grammar draws a clear distinction between (linguistic) competence and performance, and it has been declared and accepted since the earliest days of generative grammar that the object of the investigation is the competence. The model for linguistic competence is often called *Computational System*, which consists of operations generating on the basis of a numeration (i.e., a set of lexical items taken from the Lexicon[2] and formal features) a pair of abstract representations—PF and LF—that underlie the phonological representation and the semantic interpretation, respectively. One can consider the pair of PF and LF as corresponding to a so-called 'sentence'. The sentences which can be derived by this system are *grammatical*, while those which cannot are *ungrammatical*.

(1)



Thus, if the distinction between grammatical and ungrammatical sentences *were* directly observable, the investigation of competence *could* ideally be carried out as follows:

---

[1] This is basically a joint work with Hajime Hoji, but Ayumi Ueyama is responsible for adopting the particular exposition in this presentation. Hoji (in preparation) contains further empirical as well conceptual discussion of the relevant issues.

The earlier versions of this paper have been presented at the Annual Conference of Philosophy of Science Society, Japan (October, 2006) and a *Symposium on Generative Grammar and Methodology in Science*, held by Japanese Cognitive Science Society (December 2007). We are especially grateful to Yasuo Deguchi for his insightful comments and suggestions that have helped us articulate our idea in the present form. We also wish to thank Colin Phillips and Fritz Newmeyer for the intensive discussion before and during the J/K Conference 17. We have also benefited from comments and suggestions by Teruhiko Fukaya, Yutaka Kano, Kiyoko Kataoka, Emi Mukai, Yukinori Takubo, Katsuo Tamaoka, Kazuhisa Todayama and the participants in the above conferences and the lectures given at the University of Southern California, Kyushu University, Kyoto University, and Tokyo Metropolitan University.

[2] What is meant by the Lexicon is the 'mental Lexicon' of the speaker/informant where individual items are listed with their idiosyncratic properties.

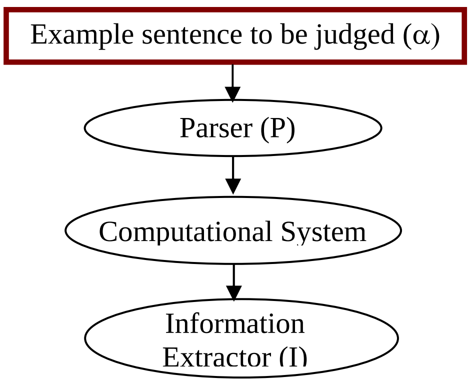
(2) 1. Identify some grammatical and ungrammatical sentences.
2. Hypothesize the Computational System so that it can derive the former but not the latter.
3. Deduce a prediction about sentences that have not yet been considered based on the hypothesized Computational System.
4. Test the prediction.
5. Proceed on the basis of the results of the test.

The primary data in actual research in generative grammar, however, is based on acceptability judgments on a given sentence (under a specified interpretation). There is a huge and in fact fundamental difference between acceptability and grammaticality. Making an acceptability judgment (under a specified interpretation) is an activity of detecting (and reporting) some sensation which is triggered by an example sentence being shown (along with the particular question posed). Grammaticality, on the other hand, is a notion having to do with whether the Computational System generates an output or not. Clarification and articulation are thus needed of the relationship between acceptability and grammaticality, especially in regard to how we can obtain data about the latter on the basis of the observation about the former. This is directly related to the testability of our hypotheses about the Computational System, without which we cannot expect to make steady progress toward the goal of discovering the properties of the Computational System.

2.2. Model of judgment making

The model of judgment making that we assume can be outlined as in (3). See Appendix for a more detailed version of the model.

(3) "How acceptable is sentence α under interpretation γ?"



Suppose that one is asked how acceptable sentence α is under the specified interpretation γ. When presented sentence α, the Parser[3], along with the word recognition, figures out which words are to form a constituent, which predicate is to take which argument(s) and so on, referring to the Lexicon when necessary. If there arises no conflict at the end of sentence α, the parsing is considered to be successful, and a numeration is formed based on the information thus obtained.[4]

---

[3] The conception of Parser adopted here is quite different from what is usually assumed in the field in regard to its function; the output of Parser as conceived here is neither an interpretation nor a syntactic structure corresponding to the presented string.

[4] Although one might object that it is rather inefficient if a speaker has to break down a sentence into words and then put them together again each time s/he tries to make an acceptability judgment on a sentence, we maintain that we *must* adopt something like (3) (or its fuller version (50) given in Appendix) as the model of judgment making once we accept the theory of the Computational System adopted here and once we commit ourselves to making our hypotheses about the Computational System empirically testable on the basis of the informants' acceptability judgments.

(4) **Parser**

a. Input : sentence α
b. Output : P(α)
c. Mechanism: Certain kind of pattern matching based on the knowledge of the language stored by the speaker through his/her linguistic experience

We assume (i) that what is available in (4c) and how easily and readily it can be utilized in the pattern matching vary depending upon each speaker within the confines imposed by the properties of the CS and what is available in the Lexicon, and (ii) that in principle P(α) does not contain sufficient information to fully determine the numeration, and hence, whatever necessary items and features are to be supplemented when an actual numeration gets formed. Let us call the numeration μ. μ is an input to the Computational System and its outputs are LF and PF representations, LF(μ) and PF(μ).

(5) **Computational System**

a. Input : Numeration μ
b. Output : LF(μ) and PF(μ)
c. Mechanism: Combination of several operations, including Merge, Move, and Agree; a completely innate system

If LF(μ) and PF(μ) obtain, it means that numeration μ yields a grammatical sentence, by definition. Notice, however, that we have to check if PF(μ) and α are non-distinct.[5]

Finally, LF(μ) goes into the Information Extractor, where it is 'interpreted into' a semantic representation, SR(μ), which is a conjunction of pieces of information conveyed by LF(μ).[6]

(6) **Information Extractor**

a. Input : LF(μ)
b. Output : SR(μ) (i.e., pieces of information conveyed by LF(μ), to be compared with γ; see (13))
c. Mechanism: Replacement of LF objects with SR objects, another innate system

Just as PF(μ) has to be compared with α, so SR(μ) has to be compared with γ. The output SR(μ) must meet the conditions specified in γ.

This is how the activity of making acceptability judgments is hypothesized here, with the Computational System embedded at its center. This activity is considered to be an act of judging α acceptable only if:

(7) a. PF(μ) is non-distinct from α, and
b. SR(μ) meets the conditions specified in γ.

2.3. An illustration

Let us review the proposed model of judgment making by the informant, this time with a concrete example. Suppose that one is asked how acceptable (8) is under the interpretation that *Toyota* and *asoko* refer to the same entity.[7]

(8) α <u>Toyota</u>-ga <u>asoko</u>-no sitauke-o uttaeta.

[5] Obviously, if they are different from each other, this activity does not qualify as a judgment of α. It is therefore necessary for a researcher to devise a way to choose only informants who are attentive enough to detect any difference between the PF and α when it arises.

[6] See (13) below, for example.

[7] Notice that γ is not the interpretation of the *entire sentence* in (8).

Toyota-NOM that-GEN subsidiary-ACC sued
'Toyota sued its subsidiary.'

(9) γ *Toyota* and *asoko* refer to the same entity.

First, the Parser works on sentence α (8) and figures out what is given in (10).

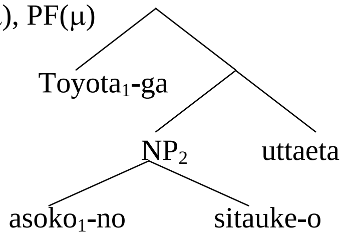
(10) P(α) *Toyota-ga* is an argument of *uttaeta* 'sued'.
asoko-no modifies *sitauke-o* 'subsidiary'.
The phrase whose head is *sitauke-o* is an argument of *uttaeta*.

Since no conflict arises at the end of sentence α, the parsing is considered to be successful, and numeration μ is formed based on the information in (10).[8]

(11) Numeration μ
{Toyota$_1$-ga, asoko$_1$-no, sitauke$_2$-o, uttaeta }

Then an LF representation such as (12) obtains, and the PF representation also looks like (12) in this case.

(12) LF(μ), PF(μ)



Since (8) is compatible with (12), this derivation satisfies the requirement in (7a).

LF(μ) further goes into the Information Extractor, and SR(μ) (13) is derived. SR(μ) is a conjunction of the four statements in (13a-d).[9]

(13) SR(μ)
a. uttaeta(x_2)(x_1)
b. x_1 : Toyota
c. x_1 : asoko
d. x_2 : sitauke(x_1)

Since (13) is compatible with what is specified as γ in (9), the judgmental sensation obtains that α is acceptable under the specified interpretation γ.

2.4. Sense of acceptability

The schema in (14) formalizes the sense of acceptability β, representing the 'full acceptability' and the

---

[8] The indexical numbers can be considered to be so-called referential indices, whose function cannot be discussed here due to space considerations. The numeration may contain more abstract features or items, but they are omitted here since they are irrelevant to the current discussion. As indicated in (11), it is assumed here that case markers are already present in the numeration.

[9] Each statement in (13a-d) can be read as follows:
(i) a. x_1 sued x_2
b. x_1 is an entity which is called 'Toyota'
c. x_1 is an entity which can be designated by *asoko*
d. x_2 is an entity which is a subsidiary of x_1

'complete unacceptability' as 1 and 0, respectively.[10]

(14) Sense of acceptability (which ranges between 0 and 1):
i) $\beta = 0$ if [G] = 0
ii) $\beta = [G] - [P] - [I] + \varepsilon_i$ if [G] = 1, where
[G] is 1 if SR(μ) compatible with γ obtains; otherwise, [G] is 0.[11]
[P] is some value (between 0 and 1) which expresses the difficulty in Parsing.
[I] is some value (between 0 and 1) which expresses the unnaturalness of SR(μ).
ε_i is random error.

According to (14), β is necessarily 0 if [G] is 0.[12] There can be several cases in which [G] is 0.

(15) [G] is 0 in any of the following cases:
a. Parsing has failed, resulting in the failure of numeration formation.
b. (Parsing has been successful, but) the derivation from μ to LF(μ) and PF(μ) has failed.
c. (Parsing has been successful and the derivation of LF(μ) and PF(μ) has been successful, but) the derivation from LF(μ) to SR(μ) has failed.
d. (Parsing has been successful, the derivation of LF(μ) and PF(μ) has been successful, and the derivation of SR(μ) has been successful, but) the SR(μ) is not compatible with γ.

Although the cases in (15a-d) are quite different from each other, they all result in β being 0. Therefore, we assume that the sense experiences for the cases in (15a-d) are in principle not distinguishable, at least on the basis of the informant judgments.

β can be greater than 0 only if [G] is 1. Therefore, as long as α is parsed with sufficient attentiveness, (16) must hold.[13]

(16) If α is grammatical, its β is some value between 0 and 1.
If α is ungrammatical, its β is always 0.

In the case of the example in (8), the value of [G] is 1, and the value of [P] should presumably be quite close to zero, since it is one of the 'basic' constructions. The value of [I] may not be zero, however. The felicitous use of expression *asoko* in Japanese requires that the user (i) know the entity by direct experience, and (ii) feel it to be 'not proximal'.[14] The value of [I] may increase if the person who judges the sentence fails to control these factors at the time of judging it. The value of [I] may decrease, however, if the same person gives it another try and successfully controls the factors in question. Thus, one of the main claims in this paper is that the acceptability value of a sentence under a specified interpretation is not necessarily something inherently determined or constant but can in principle vary a great deal depending on people as well as on occasions.

---

[10] The conception of the sense of acceptability in (14) is inspired by a suggestion made by Y. Deguchi (p.c., September 2007).

[11] It is assumed here that only those who are attentive enough as to the surface string presented and the interpretation specified in the experiment are eligible as an informant for an acceptability judgment experiment. Concrete methods of identifying 'eligible informants' will have to be addressed in a separate paper.

[12] This contrasts with the case when [G] is 1, since β need not be 1 even if [G] is 1.

[13] One might object that we sometimes take an ungrammatical sentence to be acceptable, perhaps because of overlooking something or other. But in such cases, (7a) is not satisfied. See footnote 11.

[14] This reflects a view proposed in Hoji et al. 2003, which can be regarded as an extension of Kuroda 1979 and Takubo & Kinsui 1996.

3. Hypotheses and observations

Now that the model of judgment making has been introduced, we are ready to consider how a proposal in generative grammar is to be tested empirically. Since a model consists of a set of hypotheses, it is impossible to examine the empirical adequacy of one particular hypothesis in isolation, strictly speaking. But suppose that a theory, T_0, consists of *n* hypotheses (*n* a number).

(17) $T_0 = \{H_1, H_2, ... H_{n-1}, H_n\}$

Even if we cannot evaluate H_n in isolation, we can still compare T_0 with T_1 in (18).

(18) $T_1 = \{H_1, H_2, ... H_{n-1}\}$

Or, it is also possible to compare T_0 with T_2 in (19).

(19) $T_2 = \{H_1, H_2, ... H_{n-1}, H_q\}$

In any case, as long as the research is equipped with some heuristic conventions that make it possible to evaluate theories, it is expected to make steady progress. The following illustration can be considered as a case in which T_0 is compared with T_1.

3.1. Claims, Schemata and Examples

For illustration, suppose that one is considering whether to add a hypothesis stated in (20) to the theory.

(20) Hypothesis (regarding the Computational System): Anaphor X is licensed only if there is some other element Y which satisfies all of the following conditions.
- i) Y c-commands X.
- ii) X and Y are co-arguments.
- iii) X and Y share the ϕ-features (such as gender, number, person)

(20) is a hypothesis regarding the Computational System and hence it is stated strictly in terms of theoretical concepts and relations. We thus need a 'bridging proposition' which connects the theory of the Computational System to sense experiences (i.e., the informant's intuitions), so as to be able to empirically evaluate the theory containing this hypothesis. Let us call a 'bridging proposition' in this sense a *Claim*. (21) is an instance of a Claim for the hypothesis in (20).

(21) Claim:
[... Y ... X ...], where X is an anaphor, is acceptable only if
- i) Y c-commands X,
- ii) X and Y are co-arguments, and
- iii) X and Y share the ϕ-features (such as gender, number, person)

We are not yet ready for carrying out an empirical test, however. While a Claim contains an expression referring to sense experiences (i.e., 'acceptable' in the case of (21)), it also has theoretical concepts as given in (i)-(iii) in (21), including the hierarchical notion of *c-command*. Notice that acceptability judgment is made upon the presentation of an example sentence α, accompanied by the specified interpretation γ, but α is presented to the informant only in terms of the linear relations among the elements therein. Therefore, the researcher has to make a commitment as to (i) how a linearly arranged string of words would correspond to an LF representation which contains the 'intended' structural relations among the items in question,[15] and (ii) which words should be used to

[15] Since an example sentence must go into Parser before entering the Computational System, the conversion of a Claim into Schemata most often requires assumptions about Parsing.

test the Claim in question. Let us refer to a general pattern of example sentences to be judged with a specified interpretation as a *Schema*.[16] By definition, a Schema can only describe what types of words are arranged in what order, and what kind of interpretation is at stake. For instance, the Claim in (21) contains a condition referring to a structural relation of *c-command*, which is a notion in the Computational System. Since a relation in an LF representation is not directly visible to us, we need to specify some 'construction' in which Y unambiguously c-commands X, such as one in which Y is a subject and X is an object. In addition, we also need to specify which expressions are *anaphors* in the sense of (21).

(22) Hypothesis regarding Lexicon:
A reflexive pronoun in English is an anaphor.

Thus, (23) is one Schema which corresponds to a case in which the reading in question is possible under (21).

(23) ^{ok}Schema$_{1-1}$:
[NP1 V NP2], where NP2 is a reflexive pronoun, and NP1 and NP2 share the ϕ-features, can be acceptable.

Obviously, Subject-Object is not the only case in which the intended structural relation obtains, and more Schemata could be constructed.

In addition, since (21) contains three conditions (i)-(iii), it follows that there are at least three patterns in which the reading in question is claimed to be impossible.

(24) [... Y ... X ...], where X is an anaphor, is unacceptable if
i) Y does **not** c-commands X.
ii) X and Y are co-arguments.
iii) X and Y share the ϕ-features (such as gender, number, person)

(25) [... Y ... X ...], where X is an anaphor, is unacceptable if
i) Y c-commands X.
ii) X and Y are **not** co-arguments.
iii) X and Y share the ϕ-features (such as gender, number, person)

(26) [... Y ... X ...], where X is an anaphor, is unacceptable if
i) Y c-commands X.
ii) X and Y are co-arguments.
iii) X and Y do **not** share the ϕ-features (such as gender, number, person)

Each of (24)-(26) can be converted into Schemata, just as in the case of (23), yielding (27)-(29), for example.

(27) *Schema$_{1-1}$:
[NP2 V NP1], where NP2 is a reflexive pronoun, and NP1 and NP2 share the ϕ-features, is unacceptable.

(28) *Schema$_{2-1}$:
[NP1 V [that NP V NP2]], where NP2 is a reflexive pronoun, and NP1 and NP2 share the ϕ-features, is unacceptable.

(29) *Schema$_{3-1}$:
[NP1 V NP2], where NP2 is a reflexive pronoun, and NP1 and NP2 do not share the ϕ-features, is unacceptable.

---

[16] In other words, a Schema is a bundle of instructions according to which any researcher can make example sentences relevant to the hypothesis in question.

I only showed one instance of Schema for each pattern, but obviously, additional schemata can be easily designed. Thus the entire picture will look like (30):

(30)

Hypothesis
|
Claim
|

^{ok}Schema$_{1\text{-}1}$	^{ok}Schema$_{1\text{-}2}$	^{ok}Schema$_{1\text{-}3}$
*Schema$_{1\text{-}1}$	*Schema$_{1\text{-}2}$	*Schema$_{1\text{-}3}$
*Schema$_{2\text{-}1}$	*Schema$_{2\text{-}2}$	*Schema$_{2\text{-}3}$
*Schema$_{3\text{-}1}$	*Schema$_{3\text{-}2}$	*Schema$_{3\text{-}3}$

The examples in (31)-(34) are instances of (23), (27)-(29), respectively. Again, obviously, numerous examples can be provided for each schema in (30), and eventually the whole family of examples will be built as in (35), where the empirical consequences of the Claim are specified and organized as discussed above.

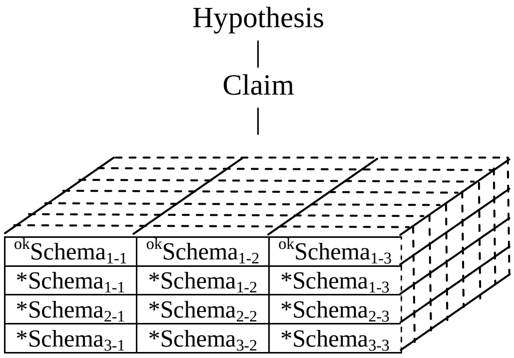
(31) ^{ok}Example$_{1\text{-}1\text{-}1}$:
John loves himself.

(32) *Example$_{1\text{-}1\text{-}1}$:
Himself loves John.

(33) *Example$_{1\text{-}2\text{-}1}$:
John thinks that Mary loves himself.

(34) *Example$_{1\text{-}3\text{-}1}$:
John loves herself.

(35)



3.2. Representative Values

The acceptability of each of the Examples behind each Schema is testable. It is therefore not unreasonable to assume that a *representative value* (**RV**) can be calculated for each Schema on the basis of **RV**s of acceptability value of Examples behind each Schema.

RV of an Example can vary depending on the factors given in (36), and **RV** of a Schema can vary according to the choice of Examples.

(36) Sources of judgment variation:
a. an informant who makes the judgment
b. other reasons that are beyond our comprehension (at the present time or even inherently)

It is therefore necessary to consider a minimally sufficient amount of data to make the **RV** statistically significant. Let us suppose (37) in order to illustrate the issue more concretely.

(37) a. There are *m* informants for this experiment.
b. Each informant is asked the acceptability value of each example sentence *q* times.
c. The number of *Examples which realize *Schema$_{1\text{-}1}$ is n.

The answers of Speaker 1 for *Ex$_{1\text{-}1\text{-}k}$ (where k is between 1 and n), for example, may not be consistent across the *q* times that Speaker 1 has judged *Ex$_{1\text{-}1\text{-}k}$. But we can take the average and consider that to be the *representative value* of *Ex$_{1\text{-}1\text{-}k}$ for Speaker 1 (i.e., **RV**(*Ex$_{1\text{-}1\text{-}k}$, Speaker 1)). The cells marked by <A> in (38) will be filled by this type of **RV**.

(38)

***Schema$_{1\text{-}1}$**	*Ex$_{1\text{-}1\text{-}1}$	*Ex$_{1\text{-}1\text{-}2}$	...	*Ex$_{1\text{-}1\text{-}n}$	
Speaker 1	<A>	<A>		<A>	<C>
Speaker 2	<A>	<A>		<A>	<C>
Speaker 3	<A>	<A>		<A>	<C>
...					
	<B>	<B>		<B>	<D>

When all the <A> cells are filled, we can calculate the <B> cells, which are the representative value of *Ex$_{1\text{-}1\text{-}k}$ across informants (i.e., **RV**(*Ex$_{1\text{-}1\text{-}k}$)). We can also calculate the <C> cells, which are the representative value of *Schema$_{1\text{-}1}$ for a certain informant (i.e., **RV**(*Schema$_{1\text{-}1}$, Speaker x)). Then <D>, the representative value of *Schema$_{1\text{-}1}$ across informants (i.e., **RV**(*Schema$_{1\text{-}1}$)), is derived, and ultimately (39) obtains, where each Schema has been assigned its **RV**.

(39)

Hypothesis
|
Claim
|

^{ok}Schema$_{1\text{-}1}$	^{ok}Schema$_{1\text{-}2}$	^{ok}Schema$_{1\text{-}3}$
*Schema$_{1\text{-}1}$	*Schema$_{1\text{-}2}$	*Schema$_{1\text{-}3}$
*Schema$_{2\text{-}1}$	*Schema$_{2\text{-}2}$	*Schema$_{2\text{-}3}$
*Schema$_{3\text{-}1}$	*Schema$_{3\text{-}2}$	*Schema$_{3\text{-}3}$

↕

RV(^{ok}Schema$_{1\text{-}1}$)	**RV**(^{ok}Schema$_{1\text{-}2}$)	**RV**(^{ok}Schema$_{1\text{-}3}$)
RV(*Schema$_{1\text{-}1}$)	**RV**(*Schema$_{1\text{-}2}$)	**RV**(*Schema$_{1\text{-}3}$)
RV(*Schema$_{2\text{-}1}$)	**RV**(*Schema$_{2\text{-}2}$)	**RV**(*Schema$_{2\text{-}3}$)
RV(*Schema$_{3\text{-}1}$)	**RV**(*Schema$_{3\text{-}2}$)	**RV**(*Schema$_{3\text{-}3}$)

This way, a given hypothesis regarding the Computational System is connected to its empirical predictions (i.e., ^{ok}Schemata and *Schemata) which can be tested empirically by calculating the **RV**s.

3.3. Empirical adequacy of a theory

If the **RV**s of ^{ok}Schemata and *Schemata are as given in (40), it is safe to conclude that the theory under examination is empirically well-motivated.

(40) Case A:

RV(^{ok}Schema$_a$)	1	**RV**(*Schema$_d$)	0
RV(^{ok}Schema$_b$)	1	**RV**(*Schema$_e$)	0
RV(^{ok}Schema$_c$)	1	**RV**(*Schema$_f$)	0
…		…	

In contrast, if the RVs are like (41), where there is no *Schema whose **RV** is 0 or even close to it, it is natural to regard the theory as not being supported empirically.

(41) Case B:

RV(^{ok}Schema$_a$)	0.6	**RV**(*Schema$_d$)	0.5
RV(^{ok}Schema$_b$)	0.5	**RV**(*Schema$_e$)	0.7
RV(^{ok}Schema$_c$)	0.7	**RV**(*Schema$_f$)	0.6
…		…	

If the addition of a new hypothesis has resulted in Case B, it should not be allowed to stay in the theory as it is, since it has yielded a prediction that is contradicted by observations.[17]

It is not always easy, however, to clearly state when a hypothesis fails. Consider Cases C-E as indicated below.

(42) Case C:

RV(^{ok}Schema$_a$)	0.6	**RV**(*Schema$_d$)	0
RV(^{ok}Schema$_b$)	0.5	**RV**(*Schema$_e$)	0
RV(^{ok}Schema$_c$)	0.7	**RV**(*Schema$_f$)	0
…		…	

(43) Case D:

RV(^{ok}Schema$_a$)	0.8	**RV**(*Schema$_d$)	0.3
RV(^{ok}Schema$_b$)	0.7	**RV**(*Schema$_e$)	0.4
RV(^{ok}Schema$_c$)	0.6	**RV**(*Schema$_f$)	0.2
…		…	

(44) Case E:

RV(^{ok}Schema$_a$)	0.6	**RV**(*Schema$_d$)	0
RV(^{ok}Schema$_b$)	0.5	**RV**(*Schema$_e$)	0.3
RV(^{ok}Schema$_c$)	0.7	**RV**(*Schema$_f$)	0
…		…	

We would like to suggest that Case C passes, Case D fails, and Case E merits a 'reexamination test', so to speak. Let us briefly discuss why.[18]

As stated in section 2.3, it is assumed here that a sentence is felt to be unacceptable not only when it is ungrammatical but also when the value of [P] and/or [I] in (14) are/is too big, or when parsing has failed.

(14) Sense of acceptability (which ranges between 0 and 1):
i) $\beta = 0$ if $[G] = 0$
ii) $\beta = [G] - [P] - [I] + \varepsilon_i$ if $[G] = 1$, where
[G] is 1 if SR(μ) compatible with γ obtains; otherwise, [G] is 0.
[P] is some value (between 0 and 1) which expresses the difficulty in Parsing.
[I] is some value (between 0 and 1) which expresses the unnaturalness of SR(μ).
ε_i is random error.

(15) [G] is 0 in any of the following cases:
a. Parsing has failed, resulting in the failure of numeration formation.

[17] One may object by pointing out that science progresses in the ocean of anomalies anyway (see works such as Lakatos 1970 and Feyerabend 1975, for example), and maintain that such a requirement as stated in the text goes directly against what is practiced in mature sciences. Hoji (in preparation) provides reasons why such an objection is misguided in the context of generative grammar at least at its present stage.

[18] We can provide below only limited discussion regarding the distinctions among different cases like these and related matters; the readers are referred to Hoji 2003 and Hoji (in preparation) for more examples and arguments.

b. (Parsing has been successful, but) the derivation from μ to LF(μ) and PF(μ) has failed.
c. (Parsing has been successful and the derivation of LF(μ) and PF(μ) has been successful, but) the derivation from LF(μ) to SR(μ) has failed.
d. (Parsing has been successful, the derivation of LF(μ) and PF(μ) has been successful, and the derivation of SR(μ) has been successful, but) the SR(μ) is not compatible with γ.

An **RV**(^{ok}Schema) not being as high as expected therefore should not count as fatal evidence against the theory of the Computational System. The reason why the particular **RV**(^{ok}Schema) has obtained may not have anything to do with properties of the Computational System under discussion.

The situation is totally different in the case of an **RV**(*Schema). If the sentence is ungrammatical, [G] is 0 (i.e., the case of (15b)) and β is necessarily 0. Therefore, the fact that an **RV**(*Schema) is not as low as zero is indeed devastating for the theory in question. It is for this reason that we suggest that Case C in (42) passes while Case D in (43) fails. One instance of the Case D type in (43) is a theory containing a hypothesis in (45).

(45) Hypothesis regarding Lexicon:
Zibunzisin in Japanese is an anaphor.[19]

According to (45), *zibunzisin* should exhibit a property similar to reflexive pronouns in English.

(46) (Allegedly)
a. John-ga zibunzisin-o semeta.
John-NOM self-ACC criticized
'John criticized himself.'

b. *Zibunzisin-ga John-o semeta.
self-NOM John-ACC criticized
'*Himself criticized John.'

c. *$John_1$-wa Mary-ga $zibunzisin_1$-o semeta to omotteiru.
John-TOP Mary-NOM self-ACC criticized COMP think
'*John thinks that Mary criticized himself.'

In fact it is reported in not a few works that (46b,c) are not acceptable, in support of (45). Even if both (46b) and (46c) are judged to be quite unacceptable, however, that will be a matter of an **RV**(*Example) not of an **RV**(*Schema). Since what is relevant in (39) is an **RV**(*Schema) (rather than an **RV**(*Example)), it is necessary to check other *Examples to obtain an **RV**(*Schema). As shown in (47), at least some of the alleged *Examples turn out to be quite acceptable and many such examples can in fact be constructed; see Hoji 2006: section 4.3, for example.

(47) *Examples (which are more or less acceptable):
a. $John_1$-wa [Mary-ga $zibunzisin_1$-ni horete iru to] omoikondeita.
John-TOP Mary-NOM self-DAT is:in:love:with COMP believed
'*John believes that Mary is in love with himself.'

b. $John_1$-wa [Mary-ga $zibunzisin_1$-o uragiru to-wa] omotteinakatta.
John-TOP Mary-NOM self-ACC betray COMP-TOP not:thought

[19] It is widely assumed in generative grammatical works that *zibunzisin* is a local anaphor, corresponding to English reflexives; see for example Tsujimura 1996: 225-228 for some references. The assumption continues to be used crucially in much of recent theoretical discussion, such as Nishigauchi 2002 and Saito 2003, among many others; see Hoji 2006ab for further discussion.

'*John did not think that Mary would betray himself.'

c. John$_1$-wa [Mary-ga zibunzisin$_1$-o suisensita to-bakari] omotteita.
John-TOP Mary-NOM self-ACC recommended COMP only thought

'*John firmly believed that Mary recommended himself.'

Thus, Japanese (45) is not empirically supported, unlike English (22).
We cannot simply adopt something like (48), however.

(48) A newly introduced hypothesis has to be discarded if there is an **RV**(*Schema) which is not zero.

According to the proposal in this paper, *Schemata and ^{ok}Schemata are intended to be realizations of the prediction(s) of a theory. It must be recognized, however, that a Schema qualifies as a 'theory-tester' only if the numeration(s) corresponding to an Example conforming to the Schema necessarily 'lead(s)' to the 'intended LF representation'. In the case of Case E in (44), one may argue that *Schema$_e$ does not qualify as a tester and hence **RV**(*Schema$_e$) should be ignored. This would amount to dismissing *Schema$_e$ as a *Schema, which would make Case E analogous to Case C. While more discussion is surely needed regarding this point, we assume that the new hypothesis that has resulted in Case E need not be abandoned as long as the researcher succeeds in showing that what is presented as *Schema$_e$ can in fact be structurally ambiguous and does not quite qualify as a *Schema, rendering Case E analogous to Case C, as noted above.[20]

4. Concluding remarks: progress in Generative Grammar

The predominant view in Generative Grammar about its methodology seems to have been that there is no need of methodology as such, insofar as each researcher is rational enough.[21] Such a view seems to have been a consequence of the fact that it was impossible to establish an explicit relation between grammaticality and acceptability under the model of grammar pursued in the 1960s. When Phrase Structure Rules were assumed, grammar was a static system that defined the properties of grammatical sentences. While the actual linguistic activity was assumed to make use of this knowledge, the relation between the grammatical knowledge and the actual linguistic activity remained too abstract to construct a theory for it, and as a result, no serious attempts were made to construct a model of judgment making that would embed the model of grammar at its center, specifying how the model of grammar interacts with the model of judgment making.

But after the emergence of the Minimalist Program, the situation has changed drastically. The

---

[20] There is a logical possibility that even in the case of (41), none of the Schemata qualifies as a 'theory tester'. Similarly, one can argue that the informant judgments on the examples in (47) do not logically lead us to abandon (45); but in that case, the researcher who wishes to maintain (45) bears the burden of presenting, *in addition to* a*Schema whose **RV** is (close to) 0, an ^{ok}Schema such that the examples in (47) can be regarded as conforming to it..

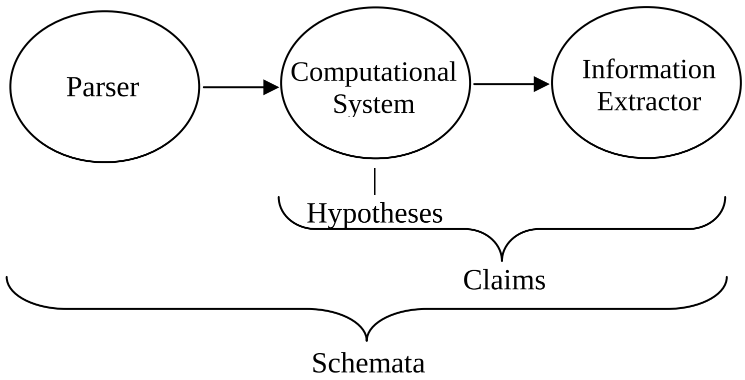
[21] See, for example, Chomsky 1988: 190. Schütze 1996 voices an opposing view; see his footnote 1 on p.210, for example. Chomsky 1977: 73 (which makes reference to the works by Kuhn and Lakatos) and Chomsky 2002: 124-128 also contain relevant remarks. Chomsky (2002: 128) states that "well-designed" [experiments] ... give the data that count, not what you come across. That's not the way linguistics was done until recently." What has been suggested in this paper, to be further articulated and illustrated with additional empirical materials in forthcoming works both by Hajime Hoji and by Ayumi Ueyama, is an attempt to spell out the basic structure of "well-designed experiments," on the basis of which we can identify "the data that count." As we elaborate in some depth elsewhere and as already noted above, the proposed view is a consequence of (i) accepting the model of the Computational System (CS) adopted here, (ii) recognizing that the informant judgment is a primary source for construction and evaluation of hypotheses about the properties of the CS, and (iii) having the desire to make progress toward the goal of discovering the properties of the CS (i.e., the desire to make generative grammar a progressive research programme in the sense of Lakatos 1970 and 1973).

most fundamental difference lies in the disposal of Phrase Structure Rules and the introduction of numeration and Merge. The Computational System can now be regarded as a dynamic system which actually builds sentences, making it possible to construct a model of actual linguistic activity with the Computational System embedded at its center.[22]

We have argued in this paper that (i) the Computational System should be considered in relation to, and more specifically as being embedded in, the theory of judgment making by the informant, and that (ii) hypotheses about the Computational System can, and in fact *should*, be connected to empirically testable observations, in terms of Claims, Schemata, and Examples. Hypotheses are confined within the Computational System. Claims relate the hypotheses to the informant's linguistic intuitions through the Information Extractor. And, finally, the construction of Schemata requires assumptions regarding parsing, in addition to the hypotheses about the Computational System and the knowledge about the Information Extractor (and also about how pragmatics may affect the informant's judgment).

The relations among Hypotheses, Claims and Schemata can be summarized as in (49).

(49)



The conception of the sense of acceptability given in (14) is an attempt to make explicit how each module contributes to or affects the informant's acceptability judgment.

What has prompted the preceding considerations is the concern for ensuring progress in Generative Grammar. If a group of Schemata exhibit the **RV** pattern such as Case A in (40) or Case C in (42), we can consider it a solid observation and hence a good candidate for an object of explanation in Generative Grammar. If the observation is solid enough, it will likely remain as an object of explanation even with a drastic change of the overall model. Thus, increasing the number of solid observations counts as empirical progress. Conceptual progress, on the other hand, obtains if we can reduce the number of theoretical concepts and relations in providing an explanation for solid observations. In either case, it is imperative that we maintain the relations among Hypotheses, Claims, Schemata, and Examples as much as possible, since it is the relations among them that would promise the status of Generative Grammar as an empirical science.

Appendix

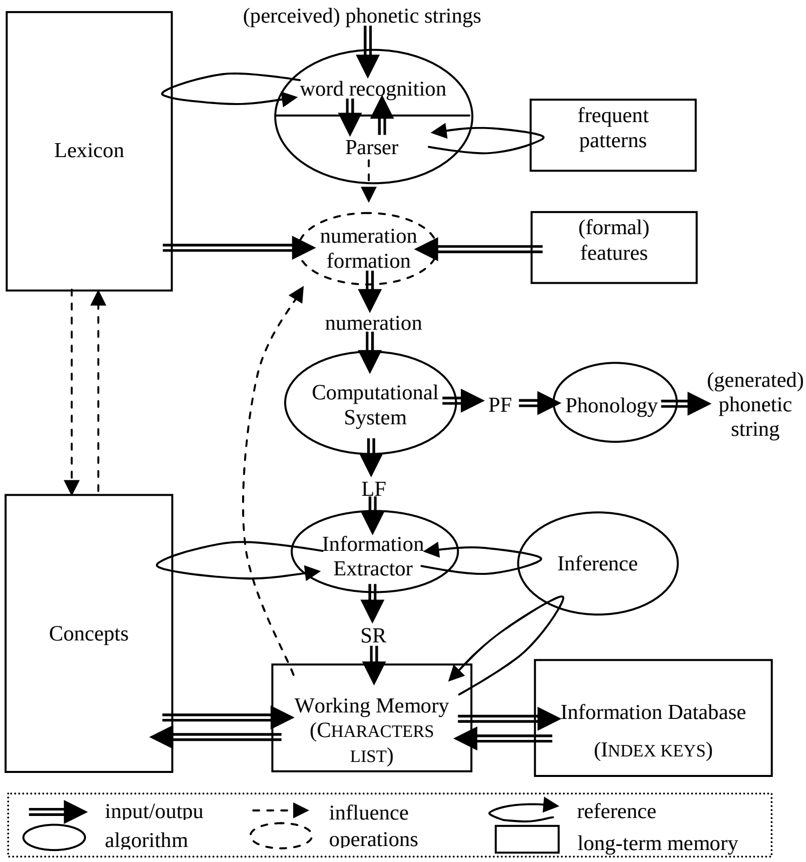
A fuller version of (3) is provided in (50).[23]

---

[22] One might still maintain that the Computational System should be considered as an abstract system that need not be related to a performance model. Such a view would however make hypotheses in Generative Grammar not empirically testable.

[23] Earlier versions of (49) are presented in Ueyama 2005, 2006, 2007, where the function of each module in (50) is addressed in some depth.

(50) Model of judgment making:



References

Chomsky, Noam (1977) *Language and Responsibility*, Pantheon Books.

Chomsky, Noam (1988) *Language and Problems of Knowledge: The Managua Lectures*, MIT Press.

Chomsky, Noam (2002) *On Nature and Language*, Cambridge University Press.

Feyerabend, Paul. (1975) *Against Method*. New Left Books. (The third edition published by Verso 1993.)

Hoji, Hajime (2003) "Falsifiability and Repeatability in Generative Grammar: A Case Study of Anaphora and Scope Dependency in Japanese," *Lingua*, vol.113, No.4-6, pp.377-446.

Hoji, Hajime (2006a) "Assessing Competing Analyses: Two Hypotheses about 'Scrambling' in Japanese," in Ayumi Ueyama, ed., *Theoretical and Empirical Studies of Reference and Anaphora—Toward the establishment of generative grammar as an empirical science*, a report of the Grant-in-Aid for Scientific Research (B), Project No. 15320052, Supported by Japan Society for the Promotion of Science, Kyushu University, pp. 139-185.

Hoji, Hajime (2006b) "Otagai," in Ayumi Ueyama, ed., *Theoretical and Empirical Studies of Reference and Anaphora—Toward the establishment of generative grammar as an empirical science*, a report of the Grant-in-Aid for Scientific Research (B), Project No. 15320052, Supported by Japan Society for the Promotion of Science, Kyushu University, pp. 126-138.

Hoji, Hajime, Satoshi Kinsui, Yukinori Takubo, & Ayumi Ueyama (2003) "Demonstratives in Modern Japanese," in A. Li & A. Simpson, eds., *Functional Structure(s), Form and Interpretation*, pp.97-128. Routledge, London.

Kuhn, Thomas. (1962) *The Structure of Scientific Revolutions*, The Chicago University Press.

Kuroda, S.-Y. (1979) "(Ko), So, A nituite (On (*ko*-), *so*- and *a*-)," in *Eigo to Nihongo to* (*English and Japanese*), Kurosio Syuppan, Tokyo, pp.41-59.

Lakatos, Imre. (1970) "Falsification and the Methodology of Scientific Research Programmes," in *Criticism and the Growth of Knowledge*, edited by Lakatos and Musgrave, 1970, Cambridge University Press. pp. 91-195.

Lakatos, Imre. (1973) "Science and Pseudoscience," a radio lecture, printed in Lakatos 1978, pp. 1-7 as "Introduction: Science and Pseudoscience," available at: http://www.lse.ac.uk/collections/lakatos//Default.htm.)

Lakatos, Imre. (1978) *The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1*, edited by John Worrall and Gregory Currie, Cambridge University Press.

Nishigauchi, Taisuke (2002) "Scrambling and Reconstruction at LF," *Gengo Kenkyu*, No. 121, pp.49-106.

Saito, Mamoru (2003) "A Derivational Approach to the Interpretation of Scrambling Chains," *Lingua*, vol.113, No.4-6, pp. 481-518.

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

Takubo, Yukinori & Satoshi Kinsui (1996) "Hukusuu no Sinteki-ryooiki niyoru Danwa-kanri (Discourse Management in terms of Multiple Mental Domains," *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, Vol.3, No. 3, pp.59-74.

Tsujimura, Natsuko, ed. (1996) *An Introduction to Japanese Linguistics*, Blackwell Publishers.

Ueyama, Ayumi (2005) "Generative Grammar as an Empirical Science---Grammaticality and Acceptability---", *Kyushu University Papers in Linguistics*, vol.25/26, pp.189-213.

Ueyama, Ayumi (2007) "Toward Generative Grammar as an Empirical Research," *Energeia: Arbeitskreis für deutsche Grammatik* 32, pp.41-55.

Ueyama, Ayumi (2008) "Information Extractor and the Condition D Effects" *BUNGAKU KENKYU (Studies in Literature)* No.105, pp.61-86, Faculty of Humanities, Kyushu University.