The Prague Bulletin of Mathematical Linguistics

NUMBER 112 APRIL 2019 55–82

Graph Theory Teaches Us Something About Grammaticality

Koji Arikawa

Department of English and Intercultural Studies, St. Andrew's (Momoyama Gakuin) University, Osaka, Japan

Abstract

Graph theory, which quantitatively measures the precise structure and complexity of any network, uncovers an optimal force balance in sentential graphs generated by the computational procedures of human natural language (C_{HL}). It provides an alternative way to evaluate grammaticality by calculating 'feature potential' of nodes and 'feature current' along edges. An optimal force balance becomes visible by expressing 'feature current' through different point sizes of lines. Graph theory provides insights into syntax and contradicts Chomsky's current proposal to discard tree notations. We propose an error minimization hypothesis for C_{HL}: a good sentential network possesses an error-free self-organized force balance. C_{HL} minimizes errors by (a) converting bottom-up flow (structure building) to top-down flow (parsing), (b) removing head projection edges, (c) preserving edges related to feature checking, (d) deleting DP-movement trajectories headed by an intermediate copy, (e) ensuring that covert wh-movement trajectories have infinitesimally small currents and conserving flow directions, and (f) robustly remedying a gap in wh-loop by using infinitesimally inexpensive wh-internally-merged (wh-IM) edge with the original flow direction.

The C_{HL} compels the sensorimotor (SM) interface to ground nodes so that Kirchhoff's current law (a fundamental balance law) is satisfied. Internal merges are built-in grounding operations at the C_{HL}–SM interface that generate loops and optimal force balance in sentential networks.

1. Introduction – Should we abandon tree notations?

For more than half a century, generative grammar, a plausible candidate for the theoretical base of biolinguistics (Chomsky, 2015), has been using tree-notation as a simple geometrical assistance to express language structures. However, Chomsky (2014) recently stated that tree notations should be abandoned because they are mis-

 Corresponding author: karikawa@andrew.ac.jp
Cite as: Koji Arikawa. Graph Theory Teaches Us Something About Grammaticality. The Prague Bulletin of Mathematical Linguistics No. 112, 2019, pp. 55–82. doi: 10.2478/pralin-2019-0002.

leading and a branching node in a tree incorrectly indicates that the node is created as a new category (ibid: at approximately 31:58).

(1) "POP [Chomsky (2013)] argues further that projection introduces no new category. That's contrary to phrase-structure grammar and all of its variants and descendants. It also follows from that that the tree notations that are commonly used are highly misleading, and probably they should be abandoned, because the reason is that there is no label for the root that branches. That's just not there. You can't avoid this in the tree notation. But it's not there. In the system, if there is no new category that is introduced by projection, it shouldn't be."

For example, when a verb V and a determiner phrase DP merge, a new set {V, DP} is created. "In its simplest terms, the Merge operation is just set formation" (Berwick and Chomsky, 2016, p. 10).

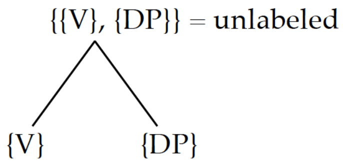

Figure 1. Before the labeling algorithm operation: Non-directed edges

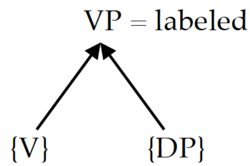

Figure 2. After labeling algorithm operation: Directed edges

Although a new "set" is created, a new "category" is not yet created, i.e., at this point, {{V}, {DP}} is unlabeled (Figure 1). The labeling algorithm (LA) given by Chomsky (2013) later identifies the nature of the set {{V}, {DP}} as the verb phrase (VP) category. Chomsky argued that a tree fails to distinguish between the pre-LA and post-LA structures. However, Chomsky's conclusions were hasty because the unlabeled merge structure before LA becomes a directed tree after LA (Figure 2).

V exists as a set that comprises subsets having phonetic features {Fphon}, semantic features {Fsem}, and formal features {Fform}, i.e., {V} = {{Fphon}, {Fsem}, {Fform}}. Similarly, a DP exists as the set {DP} = {{Fphon}, {Fsem}, {Fform}}. We refer to such a feature set as the "potential" or the "voltage" of the nodes {V} and {DP}, respectively. The merging of V and DP creates an unordered set {{V}, {DP}}, which is an unlabeled exocentric binary branching amalgam, in which the nodes {V} and {DP} are connected to the node {{V}, {DP}}. At this point, the edges are not directed, i.e., there is no feature interaction. LA identifies {{V}, {DP}} as a VP. Here, a less unified amalgam becomes a more unified compound, i.e., neither a DP nor a V. LA reduces (i.e., eliminates) a head feature [X^0] and a categorical feature [D] from {Fform} of {V} and {DP}, respectively, in the amalgam, which creates a more unified compound VP.

The feature reduction is guaranteed by the No Tampering Condition (NTC), which was deduced from the third factor principle of minimal computation (MC) (Chomsky, 2013, p. 40). Let us assume that X and Y merge, and this merger forms a new object Z. NTC specifies that neither X nor Y is modified by the Merge operation. MC requires that X and Y appear in unordered in Z (ibid). At this point, Z is an unlabeled exocentric less-unified amalgam $Z^{unlabeled} = \{\{X\}, \{Y\}\}$. When LA labels Z, a feature reduction occurs in {X} and {Y} of Z, which yields a labeled endocentric more-unified compound $Z^{labeled} = \Big\{\{\{X\} - [f_1]\}, \{\{Y\} - [f_2]\}\Big\}$, where $\{\{X\} - [f_1]\}$ and $\{\{Y\} - [f_2]\}$ indicate that formal features $[f_1]$ and $[f_2]$ are reduced from {X} and {Y}, respectively.

Such a feature reduction corresponds to a graph-theoretical "potential drop" or "voltage drop," and the potential drop drives the current flow. After Z is labeled, the linguistic features of X and Y interact with Z. We refer to such feature interactions as the "current" or the "flow." An upward feature interaction is a structure building, and a downward interaction is parsing.

In nature, currents tend to flow in the direction of energy drop, i.e., from a higher to a lower potential energy point. Thus, things fall from the points having high gravitational potentials to the points having low gravitational potentials. Similarly, steam rises from places having high energy densities (i.e., hot places) to places having low energy densities (i.e., cool places). Electric current flows from high-voltage points to low-voltage points, and air flows from areas of high atmospheric pressures to areas of low atmospheric pressures. Similarly, linguistic "current" flows (i.e., feature interaction diffuses) from the nodes V and DP having high "potential" (i.e., full set of features) to a labeled VP bearing less "potential" (i.e., having reduced or a partial set of features).

Another reason for the flow of "current" is "feature inheritance" from a strong phase head to a weak phase head, i.e., from a light verb v to a main verb V, and from a complementizer C to a tense T (Chomsky, 2008). Such a feature transportation causes a "potential drop" (i.e., feature reduction) in v and C. A flow occurs from a place of high potential to a place of low potential; therefore, feature inheritance induces a bottom-up flow. We have assumed the following properties of the structure building.

(2) *Properties of structure building*

a. *Formation of less unified exocentric set amalgam*
A syntactic object is a set comprised of a set of phonetic, semantic, and formal features. When two syntactic objects $\{\alpha\}$ and $\{\beta\}$ merge, a new set $\{\{\alpha\}, \{\beta\}\}$ is created, which is a less unified exocentric amalgam.

b. *Formation of a more unified endocentric compound*
LA makes $\{\{\alpha\},\{\beta\}\}$ an endocentric category γ, in which a formal feature [f] is reduced from $\{\alpha\}$ and $\{\beta\}$ in $\{\{\alpha\},\{\beta\}\}$, i.e., $\gamma = \{\{\alpha\}-[f_1],\{\beta\}-[f_2]\}$. This is a feature reduction.

c. *Bottom-up interaction of features*
A feature reduction caused by LA induces the upward interaction of features.

d. *Network formation*
A sequential merge followed up by LA creates a sentential network.

An LA changes an unlabeled-undirected exocentric graph to a labeled-directed endocentric network, as shown in Figure 3.

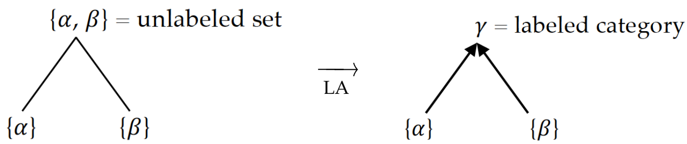

Figure 3. LA converts an unlabeled undirected exocentric graph into a labeled directed endocentric graph

The graph theory distinguishes the pre-LA and post-LA states. Contrary to Chomsky's claim, tree notions are useful for expressing sentential structures.

The remainder of this paper is organized as follows. In Section 2, we introduce the graph theory, i.e., a simple three-step version (Strang, 2016). In Section 3, we demonstrate how the graph theory can reveal a hidden force balance in simple grammatical and ungrammatical sentences. Section 4 shows that the graph theory teaches us something about the island effect. Conclusions are presented in Section 5. The Supplementary materials contain the calculation results.

2. Kirchhoff's current law governs force balance in a sentential network

We take seriously the following important tendency in nature (Strang, 2009, p. 428).

(3) Nature distributes the currents to minimize heat loss (i.e., error).

A difficult problem is what the error is relative to C_{HL}. We also assume the following general property of a network.

(4) *Properties of structure building*
A network possesses a self-organizing ability to balance the internal force in a manner such that error is minimized.

We propose the following hypothesis.

(5) *Error minimization hypothesis for C_{HL}*
A sentential network has a self-organizing ability to balance the internal force in a manner such that error is minimized.

The goal of this paper is to undertake preliminary analysis to compute the self-organizing ability of a sentential network hidden in a phrase structure and investigate whether graph-theoretical factors affect grammaticality.

The simple three-step approach is depicted as follows (Strang, 2016, p. 467). A network with nodes and edges corresponds to a network of masses and springs.

(6) *Simple three steps to uncover optimal force balance of a network*

$$\begin{tikzcd} \boxed{u} \arrow[d, "A"'] & \boxed{f} \\ \boxed{e} \arrow[r, "C"] & \boxed{y} \arrow[u, "A^T"'] \end{tikzcd} \qquad \begin{array}{ll} e = Au & A \text{ is } m \text{ by } n \\ y = Ce & C \text{ is } m \text{ by } m \\ f = A^T y & A^T \text{ is } n \text{ by } m \end{array}$$

u = Movements [potential] of n masses [nodes] = $(u_1, \cdots, u_n)$
e = Elongations [potential drop] of m springs [edges] = $(e_1, \cdots, e_n)$
y = Internal forces [current; Ohm's Law: $y = ce$] in m springs [edges] = $(y_1, \cdots, y_n)$
f = External forces [mass × gravity; KCL: $f = A^T y$] on n masses [nodes] = $(f_1, \cdots, f_n)$

Step 1 (i.e. $u \to e$) forms an incidence matrix A that expresses the geometry of a graph. Step 2 (i.e. $e \to y$) creates a conductance matrix C that measures how easily flow gets through. Ohm's Law $y = ce$ (current equals conductance times potential difference) determines a physical property c of each edge. We assign low conductance $c = 0.1$ (i.e. feature current is not easy to flow) to an XP-adjoined edge, which causes an island effect. Step 3 (i.e. $y \to f$) uses A^T (A transpose) to reveal optimal force balance hidden in the entire network, where Kirchhoff's Current Law (KCL) $A^T y = f$ (Kirchhoff (1845)) is relevant. Refer Strang (2008), Strang (2011), Strang (2016, p. 452-467) to complement this introductory section.

It is critical to point out that graph theory with KCL not only deals with the structure of an artificial object, such as an electrical circuit, it also deals with the structure of a purely mathematical and abstract geometrical graph where points are connected in various ways. We contend that graph-theoretic analysis of sentential structures is not appreciated sufficiently.

3. What does graph theory teach us about simple-sentence grammaticality?

We consider the following examples that appear to have a similar degree of structural complexity.

(7) a. He likes her.
b. * He likes she. (The intended meaning: *he* = agent, *she* = patient)

As a preliminary extension of the approach to more complex sentences, we calculate the optimal force balance hidden in island phenomena in Section 4.

3.1. Force balance hidden in a grammatical phrase structure

We demonstrate step by step how we reveal hidden force balance in a grammatical phrase structure: sample (7a). See Figure 4. The squared parts are pronounced. A viral formal feature (lower-case letters) is eliminated by its matching virus buster (uppercase letters) in a head (Piattelli-Palmarini and Uriagereka, 2004).

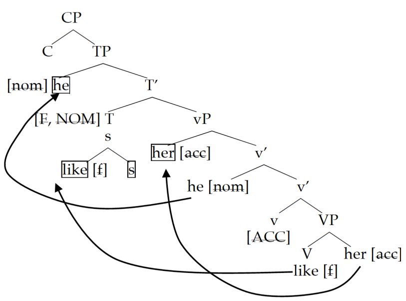

Figure 4. Grammatical phrase structure

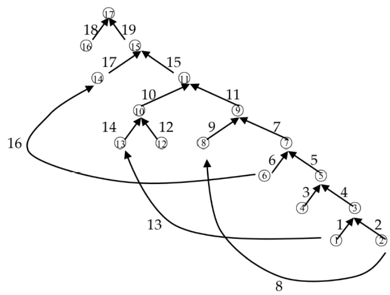

Figure 5. Graph-theoretical translation of grammatical phrase structure

We assume a set of minimal phrase-structure-building guidelines as follows.

(8) *Minimal phrase-structure-building guidelines*

a. The structure is built bottom-up.
b. The sentential heads are V, v, T, and C.
c. A set of external merge (EM; merging two terms from the structure-external Lexicon) builds a vP that contains arguments.
d. Morphological checking occurs with an internal merge (IM; structure-internal merge).
e. The sensorimotor (SM) interface externalizes one copy.

Next, we translate the single-dominance structure into a graph. See Figure 5. Assume the minimal guidelines for translating a phrase structure into a graph.

(9) *Minimal guidelines for phrase-structure-to-graph translation*

a. An IM creates a loop.
b. Nodes and edges are numbered bottom-up.
c. Matrix-clause V is numbered first.
d. Head-related nodes are numbered earlier than non-head-related nodes.
e. Head-related edges are numbered earlier than non-head-related edges.

Guideline (9a) is crucial. When node α undergoes IM, all copies of α are one and the same entity α, i.e., α appears in different places. All copies of α are related and identical. Consequently, all copies of α are connected. A loop closed by internal merge of a copy is a two-dimensional area. A graph without loops is a tree. Guideline (9b) adopts a hypothesis that structure building proceeds bottom-up. Guideline (9c) assumes that a matrix-clause predicate is the starting point of structure building. Guidelines (9d) and (9e) presuppose that a search by a virus-buster in a head is what drives IM, i.e., structural growth. The sentential graph in Figure 5 can be drawn in a plane without graph edge crossing if the graph behaves as a mobile object. KCL applies to a sentential graph because the graph is planar. A dominance relation holds in this graph-theoretic translation, and a species of the linear correspondence axiom (LCA; informally, pronounce top-down; Kayne (1994)) performs linearization in SM. Now, we translate a graph into an incidence matrix A. See Table 1.

A

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰
1	-1		1														
2		-1	1														
3				-1	1												
4			-1		1												
5					-1		1										
6						-1	1										
7							-1		1								
8		-1						1									
9								-1	1								
10										-1	1						
11									-1		1						
12										1		-1					
13	-1												1				
14										1			-1				
15											-1				1		
16						-1								1			
17														-1	1		
18																-1	1
19															-1		1

Table 1. Incidence matrix A

We use Reshish matrix calculator (RMC; matrix.reshish.com) for calculating the rank r (true size) of a matrix and for performing Gaussian elimination. For this A, $r = 16$ with computation time of 0.211s. The three rows are dependent, i.e., redundant. Rows are dependent when edges form a loop ((Strang, 2016, p. 453)) and independent when edges form a tree. C_{HL} inevitably form loops, i.e., C_{HL} leaves redundancy. Now, we transpose A to obtain a transpose matrix A^T. See Table 2.

A^T

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
①	-1												-1						
②		-1						-1											
③	1	1		-1															
④			-1																
⑤			1	1	-1														
⑥						-1										-1			
⑦					1	1	-1												
⑧								1	-1										
⑨							1		1		-1								
⑩										-1		1		1					
⑪										1	1				-1				
⑫												-1							
⑬													1	-1					
⑭																1	-1		
⑮															1		1		-1
⑯																		-1	
⑰																		1	1

Table 2. Transpose matrix A^T

We create a graph Laplacian matrix A^TA (A^T times A). See Table 3.

$A^TAx = 0$ is not invertible, i.e., not solvable. To solve an apparently unsolvable problem, typically we ground a node, i.e., make the node potential zero (Strang (2008), Strang (2011)). Grounding node ⓝ resembles hanging a spring-mass system at mass ⓝ from a ceiling. The following method is crucial to our analysis.

(10) *Ground-silent-IM-copy method for C_{HL}*
Ground a copy of IM that is not externalized at SM.

We ground IM-related nodes that are not externalized at SM, i.e., kinetic energy used for pronunciation is zero. Thus, we ground nodes ①, ②, and ⑥. The reaction force $S = s_1 + s_2 + s_3$ leaves grounded nodes and enters the root node. We obtain the network shown in Figure 6.

What are the linguistic and cognitive reasons for S? We speculate that SM contains a built-in "grounding" operation that makes at least one of IM-related copies phonetically zero. C_{HL} attempts to solve an apparently unsolvable problem by compelling SM to ground nodes, thereby calculating and creating an optimally force balanced struc-

A^TA

	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰
①	2		-1										-1				
②		2	-1					-1									
③	1	1	3		1												
④				1	-1												
⑤			-1	-1	3		-1										
⑥						2	-1							-1			
⑦					-1	-1	3		-1								
⑧		-1						2	-1								
⑨							-1	-1	3		-1						
⑩										3	-1	-1	-1				
⑪									-1	-1	3				-1		
⑫										-1		1					
⑬	1									1			2				
⑭						1								2	1		
⑮											-1			-1	3		-1
⑯																1	-1
⑰															-1	-1	2

Table 3. Graph Laplacian matrix A^TA

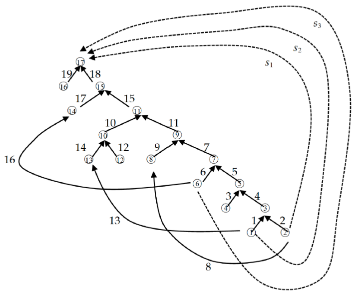

Figure 6. Reaction forces leaving grounded nodes and entering the root node

ture. SM sends it back to C_{HL}, which confirms the structural optimality and dispatches the structure with semantic features to the Conceptual-Intentional (CI) interface. C_{HL} and SM work for CI. Note that their semantic features are not zero. A remaining question is why a failure of phonetic realization in SM is sufficient to trigger grounding in C_{HL}.

Thus, nodes ①, ②, and ⑥ are reduced, i.e., they disappear from A^TA. We obtain a reduced A^TA, which we denote as A^TA_{reduced}. See Table 4.

Now, $A^TA_{\text{reduced}}x = S$ is solvable because we removed infinitely many solutions from $N(A^TA)$. RMC performs elimination and yields the following result. See Table 5.

$A^T A_{reduced}$

	③	④	⑤	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰
③	3		-1											
④		1	-1											
⑤	-1	-1	3	-1										
⑦			-1	3		-1								
⑧					2	-1								
⑨				-1	-1	3		-1						
⑩							3	-1	-1	-1				
⑪						-1	-1	3				-1		
⑫							-1		1					
⑬							-1			2				
⑭											2	-1		
⑮								-1			-1	3		-1
⑯													1	-1
⑰												-1	-1	2

Table 4. Reduced graph Laplacian matrix $A^T A_{reduced}$

	③	④	⑤	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰
③	3		-1											
④		1	-1											
⑤			5/3	-1										
⑦				12/5		-1								
⑧					2	-1								
⑨						25/12		-1						
⑩							3	-1	-1	-1				
⑪								164/75	-1/3	-1/3		-1		
⑫									101/164	-63/164		-25/164		
⑬										139/101		-25/101		
⑭											2	-1		
⑮												545/278		-1
⑯													1	-1
⑰														267/545

Table 5. Upper triangular matrix U of $A^T A_{reduced}$ after Gaussian elimination

The rank is $r = 14$. The computation time was $0.371s$. Finally, we solve the system and obtain the following result. See Table 6.

S consists of a set of syntactic features {{Fphon}, {Fsem}, {Fform}}, the potential of which is approximately equal to the total amount of node potential in TP. The result is consistent with a hypothesis that parsing is incremental (Hale, 2014). These accumulative features flow through those silent copies and return to the root node. Table 6 shows that potential is greatest in the root node ⑰ and the head C ⑯, i.e., $2.041S$, which is approximately twice that of TP ⑮, i.e., $1.041S$. A calculation reveals that the actual current of S is $S = s_1 + s_2 + s_3 = -0.999S$, which indicates that a higher node bears the cumulative potential of that of every lower node. C_{HL} recycles

Node potential	Edge current
$x_1 = 0$ (grounded)	$y_1 = -(x_3 - x_1) = -(0.022S - 0) = -0.022S$
$x_2 = 0$ (grounded)	$y_2 = -(x_3 - x_2) = -(0.022S - 0) = -0.022S$
$x_3 = 0.022S$	$y_3 = -(x_5 - x_4) = -(0.067S - 0.067S) = 0$
$x_4 = 0.067S$	$y_4 = -(x_5 - x_3) = -(0.067S - 0.022S) = -0.045S$
$x_5 = 0.067S$	$y_5 = -(x_7 - x_5) = -(0.112S - 0.067S) = -0.045S$
$x_6 = 0$ (grounded)	$y_6 = -(x_7 - x_6) = -(0.112S - 0) = -0.112S$
$x_7 = 0.112S$	$y_7 = -(x_9 - x_7) = -(0.269S - 0.112S) = -0.157S$
$x_8 = 0.135S$	$y_8 = -(x_8 - x_2) = -(0.135S - 0) = -0.135S$
$x_9 = 0.269S$	$y_9 = -(x_9 - x_8) = -(0.269S - 0.135S) = -0.134S$
$x_{10} = 0.374S$	$y_{10} = -(x_{11} - x_{10}) = -(0.561S - 0.374S) = -0.187S$
$x_{11} = 0.561S$	$y_{11} = -(x_{11} - x_9) = -(0.561S - 0.269S) = -0.292S$
$x_{12} = 0.375S$	$y_{12} = -(x_{10} - x_{12}) = -(0.374S - 0.375S) = 0.001S$
$x_{13} = 0.187S$	$y_{13} = -(x_{13} - x_1) = -(0.187S - 0) = -0.187S$
$x_{14} = 0.521S$	$y_{14} = -(x_{10} - x_{13}) = -(0.374S - 0.187S) = -0.187S$
$x_{15} = 1.041S$	$y_{15} = -(x_{15} - x_{11}) = -(1.041S - 0.561S) = -0.48S$
$x_{16} = 2.041S$	$y_{16} = -(x_{14} - x_6) = -(0.521S - 0) = -0.521S$
$x_{17} = 2.041S$	$y_{17} = -(x_{15} - x_{14}) = -(1.041S - 0.521S) = -0.52S$
	$y_{18} = -(x_{17} - x_{15}) = -(2.041S - 1.041S) = -S$
	$y_{19} = -(x_{17} - x_{16}) = -(2.041S - 2.041S) = 0$

Table 6. Node potential and edge current in the best possible force balance

potential energy (i.e., features) by compelling SM to ground silent copies. The recycled features S exit grounded nodes and enter the root node, which C_{HL} reuses for a top-down computation, i.e., parsing. An optimal force balance contains a top-down feature current.

Now we have revealed the force balance hidden in the phrase structure of *He likes her*. See Figure 7. We indicate current strength by arrow points (enlarged by a factor of 10 to make the difference among edge currents easier to see).

The sample is grammatical; therefore, C_{HL} must compute the above force balance as optimal. It is significant that current directions reverse in an optimal force balance. We speculate that structure building (the original graph) occurs bottom-up, while parsing (optimal information flow) occurs top-down. The latter corresponds to "a top down minimalist parser" that "explores a search space defined by inverting the operations of merge and move (i.e., *unmerge* and *unmove*)" (Kobele et al., 2013, p. 35). It is consistent with the statement that "grammatical categories are complex feature structures, actually calculated by the parser itself" (Hale 2014: 17). Fukui and Takano (1998) proposed a similar inverse flow, which they refer to as *demerge*, that linearizes syntactic objects top down at the SM side. We claim that such top down flows reflect a hidden self-organizing optimal force balance. C_{HL} generates an optimal force balance in which the error is minimized by eliminating two edges (i.e., edge 3, which is a head

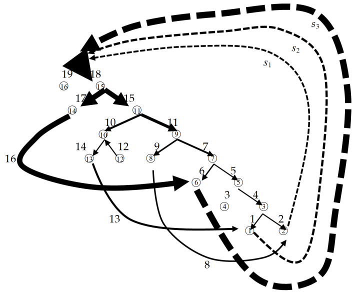

Figure 7. Hidden force balance of the grammatical sentence (7a)

projection of light verb v, and edge 19, which is a head projection of complementizer C). The optimal force balance preserves the original three independent loops. It is also significant that the current direction of edge 12 (an edge connecting two segments of V-adjoined T) is preserved.

The reaction forces $s_1(-0.521S–0.022S = -0.157S)+s_2(-0.187S–0.022S = -0.209S)+ s_3(-0.135S–0.112S = -0.633S)$ sum to $-0.999S$, which means that"gravitational force" $0.999S$ pulls the network down. Among the three IM-edges (8, 13, and 16), edge 16 ([nom]-IM edge; subject-raising trajectory) has greater resilience force ($0.521S$) than [acc]-IM-edge 8 ($0.135S$; object-raising trajectory) and [f]-IM-edge 13 (0.187S; V-raising trajectory). Edge 16 has approximately four times stronger current than that of edge 8 and roughly three times stronger current than that of edge 13. To use a spring-mass analogue, the entire network balances largely at ⑥, where the subject DP merges externally. Among the IMedges, edge 16 is analogous to a spring with the largest resilience, i.e., edge 16 works harder to adjust the balance of internal forces. In contrast, edge 8 (object-raising trajectory) is more symmetrical in that it is relatively optimal in the original graph. Node ⑥, where the subject DP merges externally, is a principal balance point of the entire network.

3.2. Force balance hidden in an ungrammatical phrase structure

For the ungrammatical sample (7b), we assume a phrase structure as in Figure 8.

Here, the [nom]-virus-checking fails. Consequently, the internal merge of *she* does not occur. We translate the phrase structure into a graph. See Figure 9.

The hidden force balance in the ungrammatical sample is as in Figure 10. Refer Supplementary 1 for the calculation.

Since the relevant sample is ungrammatical, C_{HL} must exclude the above self-organized force balance as not optimal for C_{HL}, i.e., the error is not minimized. Note that this force balance is optimal mathematically, i.e., it realizes its best possible equi-

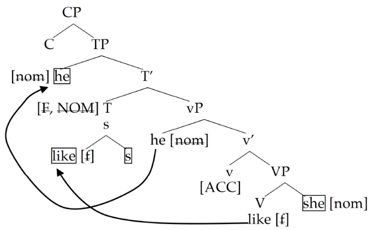

Figure 8. Phrase structure of the ungrammatical sample (7b)

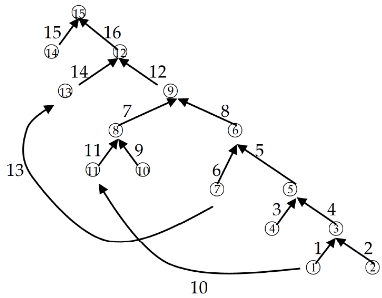

Figure 9. Graph-theoretical translation of the ungrammatical structure

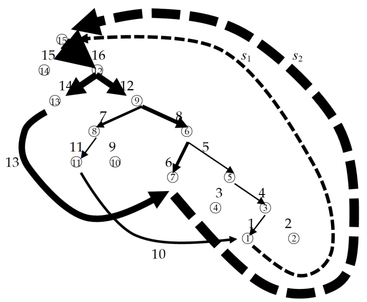

Figure 10. Hidden force balance (i.e., self-organizing ability) of the ungrammatical sample (7b)

librium and obeys KCL. However, it must contain errors that C_{HL} cannot tolerate. We consider the following as a significant observation. Unlike grammatical structure, this ungrammatical structure loses edge 2 (i.e., a complement projection of object pronoun *she*) and edge 9 (i.e., an edge connecting two segments of V-adjoined T). C_{HL} cannot tolerate the disappearance of edges 2 and 9. C_{HL} cannot delete any edge to minimize the error. We will discuss how edge disappearance contributes to grammaticality in the next section. Edge 16 (i.e., TP-to-CP projection) has the greatest resilience (–S) that pulls up the root node ⑮ to compete the "gravity." Among the two IM-edges 10 and 13, [nom]-IM edge 13 (subject-raising trajectory) has greater current force (–0.513*S*), which is approximately 3 times stronger than the other [f]-IM edge 10 (–0.18S; V-adjunction trajectory). The entire network balances principally at

	(7a) Grammatical	(7b) Ungrammatical
Number of nodes in I	17	15
Number of edges in I	19	16
Gross potential in II (S)	7.813	7.721
Absolute gross current in II (S)	4.047	3.822
Number of edge disappeared in II	2	4
— of that of I	11%	25%
Number of independent loops in I	3	2
Number of loops disappeared in II	0	0
Number of loops in II	6	4
Rank of A	16	14
Time to obtain U of A (s)	0.211	0.135
Rank of $A^T A_{\text{reduced}}$	14	13
Time to obtain U of $A^T A_{\text{reduced}}$ (s)	0.371	0.063
Absolute gross current of IM edges in II (S)	0.843	0.693
Flow direction of IM edges reversed?	Yes	Yes

(7a) *He likes her.* (grammatical) and (7b) * *He likes she.* (ungrammatical)

Table 7. Graph-theoretical properties of grammatical and ungrammatical network

7, where the subject DP merges externally. Here, edge 13 is likened to a spring with larger resilience.

3.3. Discussion—How are force balance and grammaticality related?

Here, we denote the original graph as I and the post-grounding-self-organized force balance as II. See Table 7.

A noteworthy difference between grammatical sample (7a) and ungrammatical sample (7b) is that edge 2 (complement projection) and 9 (an edge connecting two segments of V-adjoined T) submerge in sample (7b). An edge disappears when the two connecting nodes have no potential difference (i.e., potential drop), thereby no current flows along that edge. Both ends (nodes) of such an edge become disconnected. If a network loses an edge, it loses a structure and becomes more symmetrical. C_{HL} requires information flow from the complement DP for immunization of viral [acc] in the object pronoun *she*. Similarly, C_{HL} cannot tolerate loss of edge 9. A C_{HL} computation breaks down if no information flows between the two segments of V-adjoined T for immunization of viral [f] in V. Such a symmetry (no change) in the adjunction structure in its mathematically optimal balance must be an intolerable error for C_{HL}. Thus, C_{HL} must require a virus-checking operate through information flow. Both grammatical (7a) and ungrammatical (7b) lose strong-phase head (v, C). A descriptive generalization is as follows.

(11) *Descriptive generalization of force balance in simple structures*

a. C_{HL} generates an optimally-force-balanced network in which the error is minimized by disconnecting heads.

b. C_{HL} generates an optimally force balanced network in which the error is minimized by preserving edges that are related to viral formal feature checking.

Kayne (1984) was essentially correct in that a certain disconnection causes ungrammaticality. Why must heads disconnect in an optimal force balance in C_{HL}? We propose two possible answers for the puzzle.

(12) *Answer A*

When v and C merge with VP and TP, respectively, all features are transferred to V and T, respectively (feature inheritance; Chomsky (2008)). If feature inheritance precedes self-organization of force balance, no information flows from v and C when the force balance is optimal. The strong-phase-head projections from v and C must disappear to make the force balance optimal. The feature-inheritance hypothesis guarantees v = V and C = T. If v = V and C = T, V and T also submerge.

Answer B

Heads are highly symmetrical: they are in the best possible force balance in the first place. Heads are so stable and symmetric that they do not need to adjust the resilience to balance internal forces. C_{HL} uses heads as steady pivots of computation.

Putting aside which answer is preferable, observations seems to support the error minimization hypothesis for C_{HL}, i.e., a good sentential network hides a linguistically optimal force balance pattern. Our approach provides empirical evidence of the importance of current balances to grammaticality.

4. Does graph theory teach us anything about the island effect?

Here, we apply our analysis to more complex structures. We calculate force balance hidden in island-related structures (Ross, 1967), (Chomsky, 1973).

(13) *Island-effect-related examples*

a. Who$_1$ did John read [a story about t$_1$]?

b. * Who$_1$ did John read [a story that amused t$_1$]?

c. * Who$_1$ did [a story about t$_1$] amuse John?

d. John-wa [$_{DP}$ [$_{NP}$ [$_{CP}$ dare-o yorokob-ase-ta] kiji-o]] yon-da-no?
John-TOP who-ACC please-CAUSATVE-PAST article-ACC read-PAST-Q
'What is x, x a person, such that John read an article that pleased x?'

Sample (13a) indicates an overt wh-extraction from a complement DP, where no island effect is observed. Sample (13b) shows an overt wh-extraction from a complex DP that contains a relative clause CP, where an island effect is detected. Sample (13c) contains an overt wh-extraction from subject DP, where an island effect is observed. Sample (13d) is from Japanese, where the wh-phrase *dare* "who" is covertly extracted from a complex DP, as in (13b). Significantly, (13b) shows an island effect whereas (13d) does not. The wh-phrase is pronounced at the IMed position in (13b) while it is pronounced at the EMed position in (13d). For simplicity, we disregard an IM of an object with a projection of v at intermediate steps. Refer Supplementary 2 for the calculation.

4.1. Balance in overt wh-extraction from a complement DP (no island effect)

We assume the following structure for sample (13a), which is reproduced. See Figure 11.

(13) a. Who$_1$ did John read [a story about t_1]?

Unlike a pronominal complement that undergoes IM (Section 3.1), we assume that an indefinite complement DP does not undergo IM. Parentheses indicate that the term is not pronounced.

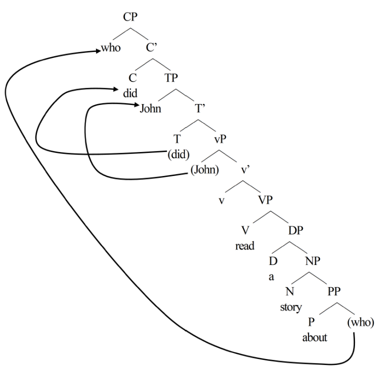

Figure 11. Overt wh-extraction from a complement DP (no island effect)

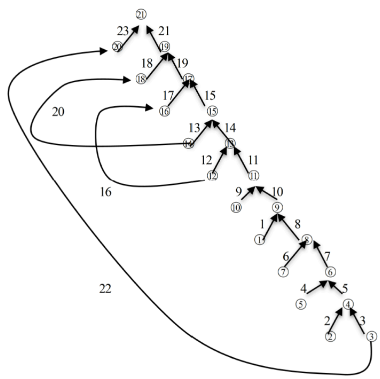

Figure 12. Graph of overt wh-extraction from a complement DP

Next, we translate the above phrase structure into the corresponding graph. See Figure 12.

The calculation reveals the following self-organizing force balance hidden in the above graph (refer Supplementary 2.1. for the calculation). See Figure 13.

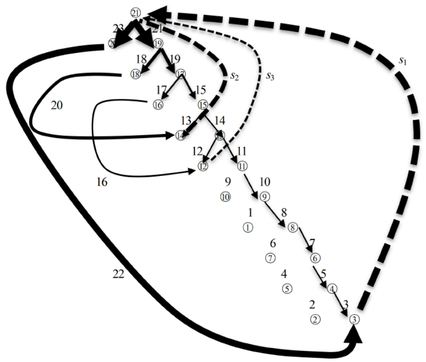

Figure 13. Balance hidden in overt wh-extraction from a complement DP (no island effect)

The optimal force balance shows top-down flow. Head edges disappear. Feature-checking-relevant edges are preserved.

4.2. Balance in overt wh-extraction from a complex DP (island effect)

We assume the following structure for sample (13b), which is reproduced. See Figure 14.

(13) b.* Who$_1$ did John read [a story that amused t$_1$]?

Why does pro not form a loop? We adopt a standard view that the feature checking of viral formal features contained in an externalized (i.e., pronounced) nominal term requires IM, which is the driving force of structural growth. We do not adopt a view in which a base-generated (i.e., externally merged) pro, which is silent, bears [nom] checked off by T by IM. Such an IM of a silent term does not contribute to the substantial structural growth. C_{HL} cannot tolerate such an unsubstantial operation; thus, pro remains at the externally-merged position, where it receives a semantic feature from v. Now we translate the phrase structure into a graph. See Figure 15.

A crucial difference between (13a) and (13b) is that the latter contains an XP-adjunction structure created by edge 13, i.e. the relative-clause CP is adjoined to the DP. Unlike head-adjunction (i.e. V-to-T head adjunction in Section 3), an XP-adjunction

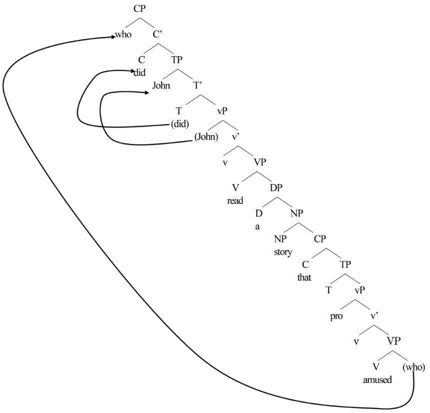

Figure 14. Overt wh-extraction from a complex DP (island effect)

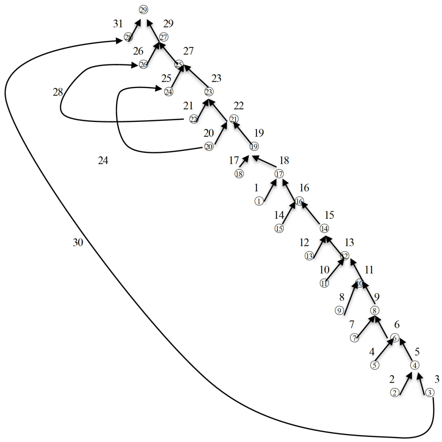

Figure 15. Graph of overt wh-extraction from a complex DP

creates an island. In particular, we assume that the conductance c of an adjoined edge is low. Let us assume that $c_{13} = 0.1$ instead of $c = 1$, which we assume for other edges. When V and NP merge, LA changes {V, NP} to VP (refer Section 1). In contrast, when CP adjoins to NP, NP embeds CP, i.e. NP contains CP. Adjoin is not Merge. If c measures edge cost as in economics (Strang, 2016, p. 458), the cost of NP-CP edge must be low because NP already contains CP. The same condition is used for calculating force balance in a Japanese example (13d) that corresponds to (13b). Refer Table 8 in Supplementary 2.2. for the reduced graph Laplacian matrix $A^T A_{\text{reduced}}$ with $c_{13} = 0.1$. The calculation (see Supplementary 2.2.) reveals the hidden force balance as in Figure 16.

Notably, a gap appears in the wh-loop, i.e. edges 15, 16, 18, and 19 that are necessary to form the wh-loop disappear in its mathematically optimal force balance. It indicates that the defective wh-loop cannot support the costly current of the wh-IM-edge 30. [1]

[1] I would like to thank an anonymous reviewer who urged me to provide a solution to an opened problem in the earlier draft, i.e. 'What causes the difference between (13a) and (13b)?', a long-standing conundrum since Ross (1967). The reviewer's request made us use a lower conductance $c = 0.1$ for an adjoined edge, which unexpectedly brought us a significant result.

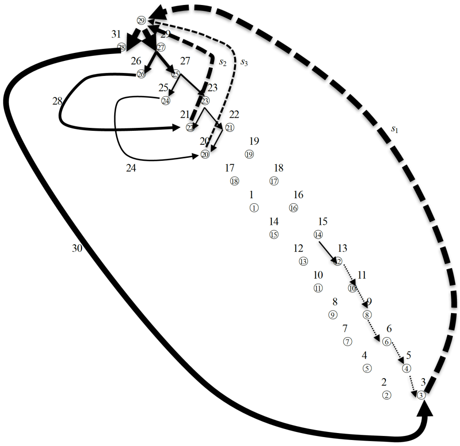

Figure 16. Balance hidden in overt wh-extraction from a complex DP (island effect)

4.3. Balance in overt wh-extraction from a subject DP (island effect)

We assume the following structure for sample (13c), which is reproduced. See Figure 17.

(13) c.* Who$_1$ did [a story about t$_1$] amuse John?

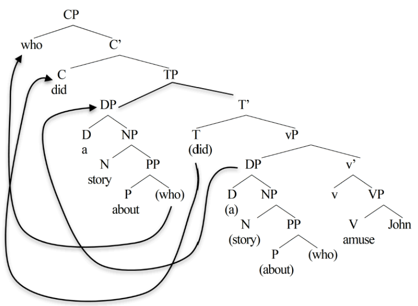

Figure 17. Overt wh-extraction from a subject DP (island effect)

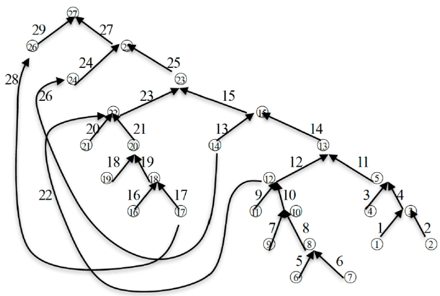

Figure 18. Graph of overt wh-extraction from a subject DP

We translate this phrase structure into a graph. See Figure 18.

The calculation (refer Supplementary 2.3.) uncovers the force balance as in Figure 19.

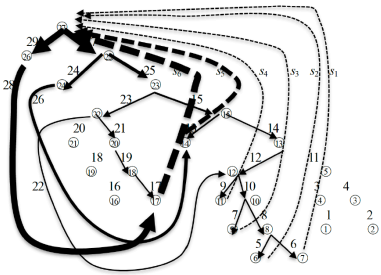

Figure 19. Balance hidden in overt wh-extraction from a subject DP (island effect)

Here, the optimal force balance shows a top-down flow. The head projection edges in the matrix clause disappear. Feature-checking-relevant edges are preserved. It is significant that head projection edges in the silent original copy of the subject island are preserved.

4.4. Balance in covert wh-extraction from a complex DP (no island effect)

We assume the following structure for a Japanese sample (13d), which is reproduced. See Figure 20.

(13) d. John-wa [$_{\text{DP}}$ [$_{\text{NP}}$ [$_{\text{CP}}$ dare-o yorokob-ase-ta] kiji-o]] yon-da-no?
John-TOP who-ACC please-CAUSATVE-PAST article-ACC read-PAST-Q
'What is x, x a person, such that John read an article that pleased x?'

We translate this into a graph. See Figure 21.

A calculation (refer Supplementary 2.4.) uncovers the hidden force balance as in Figure 22.

Remarkably, current on edge 30 (wh-IM edge) is $y_{30} = 0.0004S$, which is about 1240 times less than that of the corresponding overt wh-IM in English example (13b). Herein, the wh-IM edge preserves the original direction. Thus, it seems that such an infinitesimally small wh-IM current preserving the original flow direction does not require a complete wh-loop. Moreover, head projection edges disappear, including those in the complex-DP island. Feature-checking-relevant edges are preserved, except for edge 22, which is a DP-movement trajectory led by an intermediate copy of the topic phrase.

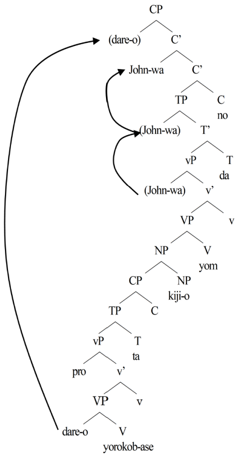

Figure 20. Covert wh-extraction from a complex DP (no island effect)

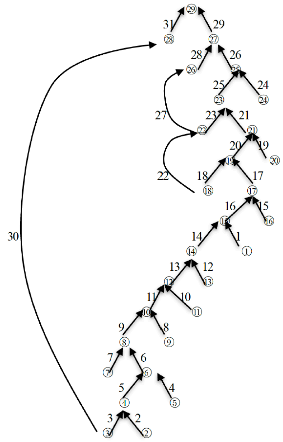

Figure 21. Graph of covert wh-extraction from a complex DP

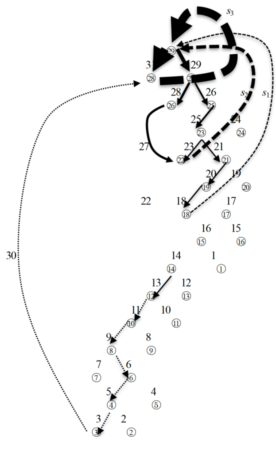

Figure 22. Balance hidden in covert wh-extraction from complex DP

4.5. Discussion—How are force balance and island effect related?

In Table 8, we highlight the properties of networks with and without the island effect. Here, the original graph and the network with self-organized force balance are abbreviated as I and II, respectively.

	(13a)	(13b)*	(13c)*	(13d)
Number of nodes in I	21	29	27	29
Number of edges in I	23	31	29	31
Gross potential in II (S)	2.933	3.157	3.181	2.141
Absolute gross current in II (S)	2.847	2.854	2.875	1.868
Number of edges disappeared in II	5	13	8	14
— of that of I	22%	42%	28%	45%
Number of independent loops in I	3	3	3	3
Number of loops disappeared in II	0	1	0	2
Number of loops in II	6	5	9	3
Rank of A	20	28	26	28
Time to obtain U of A (s)	0.419	1.029	0.36	0.138
Rank of A^TA_{reduced}	18	26	21	26
Time to obtain U of A^TA_{reduced} (s)	0.1	0.383	0.301	0.345
Absolute current of wh-IMed edges in II (S)	0.497	0.497	0.495	0.0004
Flow direction of wh-IMed edge in II reversed?	Yes	Yes	Yes	No
Does wh-loop contain adjunction structure?	No	Yes	No	Yes

(13a): grammatical overt wh-extraction from complement DP; (13b)*: ungrammatical overt wh-extraction from complex DP; (13c)*: ungrammatical overt wh-extraction from subject DP; (13d): grammatical covert wh-extraction from complex DP

Table 8. Graph theoretical properties of island-effect-related force balance

4.5.1. Grammatical (13a) versus ungrammatical (13b)*

As an anonymous reviewer pointed out, the fact that the absolute current of wh-IMed edge in II for grammatical (13a; Figure 13) and ungrammatical (13b; Figure 16)* is identical seems to indicate that our analysis fails here. However, there is a fundamental difference between the two, i.e. (13b)* lacks a wh-loop. A crucial difference between (13a) and (13b)* is that the latter contains an XP-adjunction structure (i.e. the relative-clause CP adjoins to the DP) in the wh-loop. An important condition is that an adjoined edge bears low conductance, i.e. $c_{13} = 0.1$. Therefore, the ungrammatical structure (13b; Figure 16)* has an incomplete wh-loop with a gap. No edge means no potential difference and no current flow. An incomplete wh-loop cannot support the costly wh-IM edge bearing relatively high current ($0.497S$) that reversely flows into the original wh-copy, a position to which a semantic feature is assigned. Note that C_{HL} allows an adjunction structure itself. A non-wh-sentence containing an adjoined edge is grammatical (e.g. *'John read a story that amused Mary.'*). A calculation reveals that all EM (externally-merged)-edges unrelated to loops disappear in a tree (structure without loops), i.e. they are optimal in the first place. However, C_{HL} disallows a sentential structure constructed exclusively by EM, i.e. IM must operate in C_{HL}.

4.5.2. Ungrammatical (13b)* versus grammatical (13d)

The current difference regarding the wh-IMed edge (wh-movement trajectory) between (13b)* and (13d) is remarkable. The wh-IM current of edge 30 (wh-movement trajectory) of (13b; Figure 16)* is ~1240 times greater than that of edge 30 of (13d; Figure 22). The resilience of edge 30 in (13d; Figure 22) is extremely small ($0.0004S$; relatively close to zero) and preserves the original flow direction that guarantees wh-interpretation. The same graph-theoretical result must be realized in other "wh-in-situ" languages, such as Chinese and Korean, where a similar immunity to island effect has been observed since Huang (1982). Significantly, the wh-IM edge 30 bearing infinitesimally small current and the original direction robustly remedies a gap in a wh-loop.

For C_{HL}, (13d) is grammatical because the error (i.e., "heat loss" in wh-movement trajectory) is minimized, while (13b)* is ungrammatical because the error is not minimized. A zero-current edge is likened to an inelastic wire and is symmetrical in that it is optimal in the original graph in the first place. A similar property is found in zero-current edges growing from heads (refer Section 3). It is significant that a movement trajectory of a wh-phrase that is externalized at the original position in II behaves as a head projection edge. An extremely low cost of a wh-IMed edge with the original direction is sufficient to self-balance the entire network in wh-in-situ languages. In such languages, the cost of wh-IM (wh-movement trajectory) must be very small, which Huang (1982) predicted and observed.

Huang hypothesized that the wh-IM in wh-in-situ languages takes place after spell-out (SO, i.e. a derivational point where information is sent to SM and CI). IM after SO does not affect pronunciation, thereby ensuring zero externalization cost. However, such a hypothesis faces a problem relative to why wh-IM takes place before SO in some languages (e.g. English) and after in others (e.g. Japanese). We argue against such a wh-movement parameter. In contrast to Huang's take, we assume that wh-IM (wh-movement) takes place before SO in all languages, i.e., the structure building is the same for C_{HL} of "Homosapiensese," i.e., human natural language. It is a mathematical (linear algebraic/graph theoretical) distinction of hidden force balance that causes the contrast (13b)* vs (13d). If a current is fundamentally an error, thereby causing a heat, the relevant error is minimized to a greater degree in the network of (13d; Figure 22). More specifically, the gross potential of (13b; Figure 16)* is approximately 1.4 times greater than that of (13d; Figure 22), and the absolute gross current of (13b; Figure 16)* is roughly 1.5 times stronger than that of (13d; Figure 22).

Furthermore, the current direction of the wh-IMed edge is preserved when the wh-phrase is externalized in the original position in (13d; Figure 22), unlike (13a; Figure 13), (13b; Figure 16)*, and (13c; Figure 19)*, where the wh-phrase is externalized at a higher IMed position. Thus, wh-IM in (13d; Figure 22) is also more symmetrical relative to the direction of the information flow. It is also significant that feature-checking-relevant edges are preserved, with the exception of edge 22, which is a DP-movement trajectory that is led by an intermediate copy of the topic phrase. This comprises empirical evidence that a movement trajectory between an original copy and an intermediate copy is optimal throughout the derivation, i.e., it does not need to adjust the resilience. The principal balance point of the network in (13d; Figure 22) is 28, which is the target of [wh]-checking and is the closest to the root node CP. The above observations comprise evidence for the error minimization hypothesis for C_{HL}, i.e., the force balance and current (error) minimization within the entire network affects grammaticality.

4.5.3. Grammatical (13a) versus ungrammatical (13c)*

It is significant that head projection edges in the silent original copy of the subject island are preserved in (13c; Figure 19)*. C_{HL} cannot tolerate such a head-projection-edge preservation and computes that the error is not minimized. Unlike grammatical force balance in (13a; Figure 13) and (13d; Figure 22), where the balance point is either the bottom or top of the entire network, ungrammatical (13c; Figure 16)* has their balance point at an intermediate wh-copy that is neither assigned a semantic role nor is its viral formal feature checked off. Such an ontologically weak status disqualifies an intermediate copy as an optimal balance point of the entire network. These constitute additional factors that control the error minimization hypothesis for C_{HL}. Furthermore, ungrammatical (13c; Figure 16)* hides a force balance that resembles that of the ungrammatical simple sentence **He likes she* (Section 3), where the complement *she* is disconnected from the entire structure. In other words, the terms in matrix-clause v′ are disconnected from the entire structure in (13c; Figure 16)*. A certain disconnection causes grammaticality (Kayne, 1984).

4.5.4. Simple sentence versus complex sentence

One may predict that the gross potential and absolute gross current in II of island-effect-related samples must be greater than those of simple samples because the former appear to require more energy to compute more complex structures. However, this prediction fails. As Tables 7 and 8 indicate, the net potential and absolute net current in II of simple examples are greater than those of island-effect-related examples. For C_{HL}, a simple sentence is not so simple, and a complex sentence is not so complex.

4.5.5. Why does C_{HL} contain IM?

Given the above results, we see a hint relative to answering a difficult problem, i.e., why does C_{HL} contain IM? Chomsky states that we should allow ourselves to be more puzzled as to why this is so.

(14) "Displacement [IM] had always seemed—to me in particular—a curious imperfection of language. … Pursuit of SMT [strong minimalist thesis] reveals that displacement with this property of multiple interpretation ("the copy theory of movement") is the simplest case. … This is a significant discovery, I think—too long in coming, and insufficiently appreciated, as are its consequences" (Chomsky, 2015, p. x).

SMT states that the faculty of language (FL = C_{HL}) is a perfect solution to the legibility problems that the two external interfaces (i.e., the conceptualintentional (CI) and sensorimotor (SM)) impose on C_{HL}. Consider the following example with the two copies, where the lower copy is silent.

(15) Which book did John read (which book)?
'For which x, x a book, such that John read x?'

At the initial step, verb V assigns a semantic role [patient] to the original copy of *which book* (i.e., the lower variable x) when the copy EMs with V. At a later step, C IMs with *which book* (i.e., the higher wh-phrase working as the operator binding the variable x) and the sentence is interpreted as a direct wh-question in CI. Here, MC requires one copy to be externalized. The higher copy is externalized in English-type languages, whereas the lower copy is externalized in Chinese-type languages. SMT reveals that IM is the simplest possible solution to the legibility conditions that CI and SM impose on C_{HL}. Thus, Chomsky's answer is as follow.

(16) *Why did nature create IM in C_{HL}?*
Nature created IM in C_{HL} because IM was the simplest way to balance multiple interpretation in a sentence. (Chomsky's answer)

In this paper, we add a graph-theoretic reason as to why C_{HL} contains IM, noting that IM creates loops. A crucial question to ask at this point is as follows.

(17) Do we require loops for interpretability of *any* syntactic structure? If we do, there must be loops in a sentential structure. This has thick implications for syntax.

Suppose that the following assumptions hold.

(18) a. A sentential-structure building uses an IM.

b. An IM creates a loop.
c. A sentential structure is a graph generated by C_{HL}.
d. A graph possesses balance that obey KCL to equilibrate the internal force.
e. Loops are solutions to KCL (Strang (2009), Strang (2011), Strang (2016)).

Graph theory, which is an application of linear algebra, standardly maintains assumptions (18d) and (18e). The minimalist program assumes (18a). If (18b) and (18c) hold, which is a perspective that is contra-Chomsky (2014), a sentential structure must contain loops to balance the internal force. However, more loops do not mean more optimal force balance. If C_{HL} tolerates and interprets within a certain threshold of a force-balance state, and loops are solutions to KCL, C_{HL} must require a certain pattern of force balance containing loops for interpretability in any syntactic structure. Specifically, an unpronounced IM-copy, whose phonetic externalization is determined to zero by SM to answer the legibility problems posed by C_{HL}, corresponds to a grounded node in a graph necessary for solving an apparently unsolvable problem. IM may be a built-in grounding operation that nature has created in the C_{HL}-SM interface. Information (i.e., linguistic features) flows around in a sentential network. C_{HL} needs IM to optimally self-balance the internal force in a sentential network. "What are the actual solutions to [KCL] $A^T y = 0$? The currents must balance themselves. The easiest way is to flow around a loop" (Strang, 2016, p. 456). IM may have emerged in C_{HL} because IM was the easiest way to balance currents in a sentential network. We answer Chomsky's puzzle as follows.

(19) *Why did nature create IM in C_{HL}?*
Nature created IM in C_{HL} because IM was the easiest way to balance currents and minimize errors in a sentential network. (Our answer)

5. Conclusions

In structure building, when a union set is labelled by LA (Chomsky (2013)), edges become directed, i.e., features flow upward. Contra Chomsky (2014), who claims that we should abandon graph notations in C_{HL} research, we claim that we must maintain graph notations. A graph theory equipped with KCL provides insight into grammaticality.

A significant concept that we adopt is "nature distributes the currents to minimize the heat loss (i.e., error)" (Strang, 2009). A sentential network generated by a natural object C_{HL} minimizes the error, which corresponds to what SMT refers to as a perfect solution to the legibility problems. Thus, we propose the error minimization hypothesis for C_{HL}: a good sentential network IM creates possesses a self-organizing ability to balance the internal force in a manner such that error is minimized.

We adopt Strang's simple-three-step approach of graph theory to uncover a hidden force balance in any network. Step 1 is a "geometry" step, where we translate a sentential graph (translated from a phrase structure) into an incidence matrix A. Step

2 is a "physics" step, where we investigate edge conductance matrix C. We assume that C is the identity matrix unless an edge involves XP-adjunction structure, in which case we assume $c = 0.1$. Step 3 is a "balance" step, where we use KCL $A^T y = f$ to uncover a hidden force balance in a sentential network. Here, the relevant matrix is $A^T A$ (a graph Laplacian matrix), which appears in various areas of mathematics relative to error minimization.

We calculated the hidden force balance in simple and island-effect-related sentences that are both grammatical and ungrammatical. C_{HL} minimizes errors by (a) converting bottom-up flow (structure building) to top-down flow (parsing), (b) removing head projection edges, (c) preserving edges related to feature checking, (d) deleting DP-movement trajectories headed by an intermediate copy, (e) ensuring that covert wh-movement trajectories have infinitesimally small currents and conserving flow directions, and (f) robustly remedying a gap in wh-loop by using infinitesimally inexpensive wh-internally-merged (wh-IM) edge with the original flow direction. The C_{HL} compels the sensorimotor (SM) interface to ground nodes such that Kirchhoff's current law (a fundamental balance law) is satisfied. Internal merges are built-in grounding operations at the C_{HL}-SM interface that generate loops and optimal force balance in sentential networks.

Acknowledgements

I would like to thank two anonymous reviewers for their valuable comments. The first reviewer urged me to provide a solution to an opened problem in the earlier draft, i.e., What are the graph-theoretical factors that distinguish a non-island property of a complement (non-complex) DP and island property of a complex DP containing a relative-clause CP? As hinted in the previous version, I used a lower conductance $c = 0.1$ for an adjoined edge in the latter and unexpectedly obtained a significant result. The first reviewer led us to this crucial factor. The second reviewer anticipated that the Chomskian graph-theoretical techniques are difficult to understand for people working with syntactico-semantic techniques based on lexicalized and dependency tools. However, Chomsky and his colleagues after (Chomsky, 2013) proposed to banish graph theoretical approach (e.g. (Kayne, 1984)) from the research field of language system (C_{HL}) within the minimalist (biolinguistic) program. I believe I can convince the readers that graph theory is a valuable tool to uncover the precise syntactic structure of a sentence. I would like to thank Enago (www.enago.jp) for the English language review. Finally, I would like to express my gratitude to Martin Popel, Hiroaki Sano, and Tomokazu Sano for spending their valuable time to convert my Word files to the LaTeXformat, which I am a novice at.

Bibliography

Berwick, C. Robert and Noam Chomsky. *Why Only Us: Language and Evolution*. MIT Press, 2016.

Chomsky, Noam. Conditions on Transformations. *A Festschrift for Morris Halle*, 1973.

Chomsky, Noam. On phases. *Current Studies in Linguistics Series*, 45:133, 2008.

Chomsky, Noam. Problems of projection. *Lingua*, 130:33–49, 2013.

Chomsky, Noam. Problems of Projection: Extensions, 2014. URL `https://www.youtube.com/watch?v=Icv_sCsIu6A`. A lecture at Olomouc Linguistics Colloquium.

Chomsky, Noam. *The minimalist program*. MIT press, 2015. 20th anniversary edition.

Fukui, Naoki and Yuji Takano. symmetry in syntax:merege and demerge. *Journal of East Asian Linguistics*, 7:27–86, 1998.

Hale, T. John. *Automaton Theories of Human Sentence Comprehension*. CSLI Publications, 2014.

Huang, C.-T. James. *Logical relations in Chinese and the theory of grammar*. PhD thesis, MIT, 1982.

Kayne, S. Richard. *Connectedness and Binary Branching*. Foris, 1984.

Kayne, S. Richard. *The Antisymmetry of Syntax*. MIT Press, 1994.

Kirchhoff, (von Studiosus. Ueber den Durchgang eines elektrischen Stromes durch eine Ebene, insbesondere durch eine Kreisförmige [On the transit of an electric current through a plane, in particular through a circular]. *Annalen der Physik und Chemie*, 64:487–514, 1845.

Kobele, M. Gregory, Sabrina Gerth, and John Hale. Memory resource allocation in top-down minimalist parsing. *Formal Grammar: 17th and 18th international conferences*, pages 32–51, 2013.

Piattelli-Palmarini, Massimo and Juan Uriagereka. The immune syntax: the evolution of the language virus. *Variation and Universals in Biolinguistics*, pages 341–377, 2004.

Ross, John Robert. *Constraints on variables in syntax*. PhD thesis, MIT, 1967.

Strang, Gilbert. Computational Science and Engineering I. Lecture 12: Graphs and Networks, Lecture 13: Kirchhoff's Current Law, 2008. URL `https://ocw.mit.edu/courses/mathematics/18-085-computational-science-and-engineering-i-fall-2008/index.htm`. Opencourseware. Cambridge, Massachusetts: MIT.

Strang, Gilbert. *Introduction to Linear Algebra, Fourth Edition*. Wellesley-Cambridge Press, 2009.

Strang, Gilbert. Linear Algebra. Lecture #12: Graphs, Networks, Incidence Matrices, 2011. URL `https://ocw.mit.edu/courses/mathematics/18-06sc-linear-algebra-fall-2011/`. Opencourseware. Cambridge, Massachusetts: MIT.

Strang, Gilbert. *Introduction to Linear Algebra, Fifth Edition*. Wellesley-Cambridge Press, 2016.

Address for correspondence:
Koji Arikawa
`karikawa@andrew.ac.jp`
Department of English and Intercultural Studies,
St. Andrew's University (Momoyama Gakuin)
1-1, Manabino, Izumi, Osaka, 594-1198