BALE 2008 @ York

Recursion, Modularity and the *Evo-Devo* of Language

Koji Fujita
Kyoto University

"To create is to recombine." - F. Jacob

"… an evolutionary novelty may result from the combination of two pre-existing parts with unrelated functions." - M. Ridley

"Evolution has recruited for language purposes brains structures that performed other functions in non-human primates." - T. Deacon

"… domain-specificity of language is reduced to some special arrangement of elements that are not language-specific." - N. Chomsky

Generative Biolinguistics

- Human Nature and Language Organ

(1) Design …….. Microgenesis
(2) Development … Ontogenesis
(3) Evolution ……. Phylogenesis

(1) Descriptive Adequacy ➡ <PHON,SEM>

⬍

(2) Explanatory Adequacy ➡ I-Language

⬍

(3) Evolutionary Adequacy ➡ UG

- Logical Problem of Language Evolution
- Logical Problem of Language Acquisition

Biological Evolution and Language Evolution

- Language evolution is an instance of biological evolution (in addition to cultural evolution).

→ If one's theory of biological evolution fails to account for the evolution of language, then it needs a serious reconsideration.

- Arrival of the Fittest
- Survival of the Fittest

Neo-Darwinism (Modern Synthesis)

- Adaptationist Program
  - Functionalism
- Natural Selection / Sexual Selection as the First Resort
- Gradualism

Neo-Neo-Darwinism (Expanded Synthesis)

- Non-adaptationist Program
  - Formalism
- Pluralism
  - NS/SS as the Last Resort
- Punctuated Equilibrium (saltationism?)
- Exaptation

Sexual Selection (Handicap Principle)

Figure 2. Peacock with a highly ornamented tail which, like the male quetzal's tail, evolved by female choice. If some "eyes" are removed from his tail, he becomes less attractive to peahens. It is hard to imagine how such an enormous encumbrance would be compatible with escape from predators, and indeed further enlargement of the tail may have been constrained by natural selection.

G. D. Dimijian: Evolution of sexuality: biology and behavior.

Aptation	Adaptation	NS shapes the character for a current use.
	Exaptation	A character, previously shaped by NS for a particular function (an adaptation), is co-opted for a new use.
		A character whose origin cannot be ascribed to the direct action of natural selection (a non-aptation), is co-opted for a current use.

Gould and Vrba 1982.

Adaptationism: Three Kinds

(1) Empirical Adaptationism
(2) Explanatory Adaptationism
  a. Weak
  b. Moderate
  c. Strong
(3) Methodological Adaptationism

P. Godfrey-Smith 2001; T. Shanahan 2004.

Original Function, Current Utility

- Language for Thought (internalization) or Communication (externalization)?
- The core computational system of human language is maladapted for communicative purposes.

(1) John$_i$ saw himself$_i$/*him$_i$.
(2) John$_i$ thinks [Mary saw *himself$_i$/him$_i$].

If Condition A effect in (1) is functionally explained, why is the situation reversed in (2)?

- Is recursion functional?

(1) The daughter of [John]'s son is [the son of John]'s daughter.

(2) The mouse [the cat [the dog [the boy owned] loved] admired] danced.

- Syntax is optimally designed to satisfy the CI interface system, not the SM system.

(1) What did you eat?
CI: [what did you eat what]
SM: [what did you eat ___]

- The functions of the components that jointly constituted the language faculty later in the hominin evolution may have had nothing to do with the current (or even original) function(s) of language.
- Animal communication may have only an indirect bearing on language evolution.

Paradox of Adaptive Selection

- In order to be adaptive as a communicative tool, language has to be already shared among individuals.

cf. mother-child bond, social grooming, etc.

- Language as a communicative tool is itself an instance of exaptation.
- "Humans use language for communication, but it may well be that the most important aspect of language is that it is used for internal representation in the brain."
  - J. Maynard Smith and E. Szathmáry

Strong Minimalist Thesis (SMT)

- Language is an optimal solution to legibility conditions.
- Unexplained elements of UG are zero.
- There is virtually nothing special about the origin and evolution of language.

The Third Factor and Teleomatic Explanation

- (Apparent) Goal-Directedness:
  - Teleological explanation
  - Teleonomic explanation
  - Teleomatic explanation

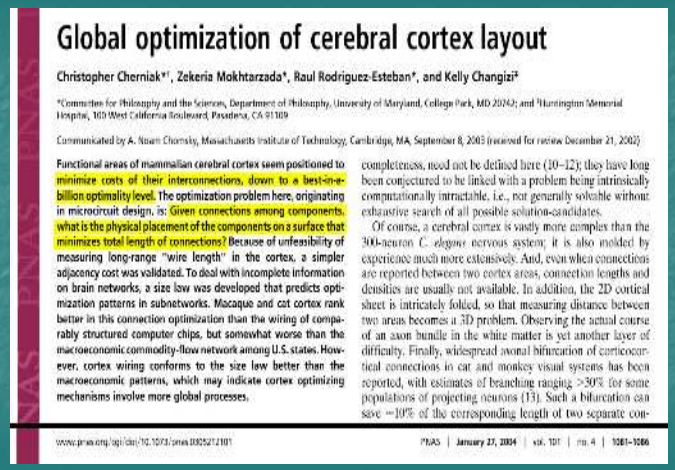
E. Mayr: *Toward a New Philosophy of Biology.*

Teleonomic process: A process of behavior that owes its goal-directedness to the operation of a program.

Teleomatic process: A seemingly end-directed process that is strictly controlled by natural laws such as the law of gravity or the first law of thermodynamics.

E. Mayr: *One Long Argument.*

The Minimalist Program is an attempt to seek a teleomatic explanation of the language design.



Global optimization of cerebral cortex layout

Christopher Cherniak*†, Zekeria Mokhtarzada*, Raul Rodriguez-Esteban*, and Kelly Changizi‡

*Committee for Philosophy and the Sciences, Department of Philosophy, University of Maryland, College Park, MD 20742; and ‡Huntington Memorial Hospital, 100 West California Boulevard, Pasadena, CA 91109

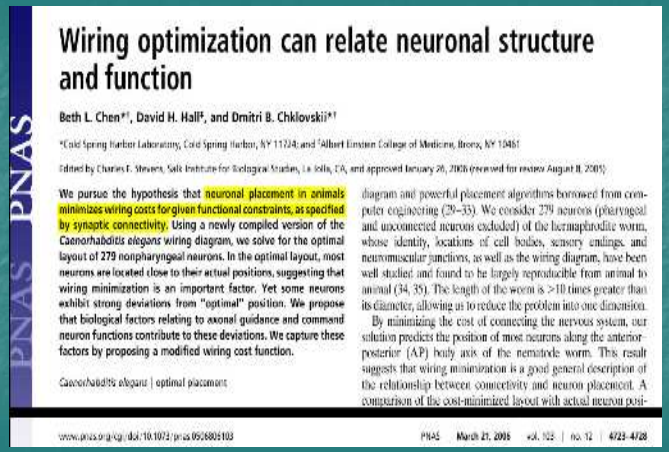
Communicated by A. Noam Chomsky, Massachusetts Institute of Technology, Cambridge, MA, September 8, 2003 (received for review December 21, 2002)

Functional areas of mammalian cerebral cortex seem positioned to minimize costs of their interconnections, down to a best-in-a-billion optimality level. The optimization problem here, originating in microcircuit design, is: Given connections among components, what is the physical placement of the components on a surface that minimizes total length of connections? Because of unfeasibility of measuring long-range "wire length" in the cortex, a simpler adjacency cost was validated. To deal with incomplete information on brain networks, a size law was developed that predicts optimization patterns in subnetworks. Macaque and cat cortex rank better in this connection optimization than the wiring of comparably structured computer chips, but somewhat worse than the macroeconomic commodity-flow network among U.S. states. However, cortex wiring conforms to the size law better than the macroeconomic patterns, which may indicate cortex optimizing mechanisms involve more global processes.

completeness, need not be defined here (10–12); they have long been conjectured to be linked with a problem being intrinsically computationally intractable, i.e., not generally solvable without exhaustive search of all possible solution-candidates.

Of course, a cerebral cortex is vastly more complex than the 300-neuron *C. elegans* nervous system; it is also molded by experience much more extensively. And, even when connections are reported between two cortex areas, connection lengths and densities are usually not available. In addition, the 2D cortical sheet is intricately folded, so that measuring distance between two areas becomes a 3D problem. Observing the actual course of an axon bundle in the white matter is yet another layer of difficulty. Finally, widespread axonal bifurcation of corticocortical connections in cat and monkey visual systems has been reported, with estimates of branching ranging >30% for some populations of projecting neurons (13). Such a bifurcation can save ≈10% of the corresponding length of two separate con-

www.pnas.org/cgi/doi/10.1073/pnas.0305271101 PNAS | January 27, 2004 | vol. 101 | no. 4 | 1081–1086



PNAS

Wiring optimization can relate neuronal structure and function

Beth L. Chen*†, David H. Hall‡, and Dmitri B. Chklovskii*†

*Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; and ‡Albert Einstein College of Medicine, Bronx, NY 10461

Edited by Charles F. Stevens, Salk Institute for Biological Studies, La Jolla, CA, and approved January 26, 2006 (received for review August 8, 2005)

We pursue the hypothesis that neuronal placement in animals minimizes wiring costs for given functional constraints, as specified by synaptic connectivity. Using a newly compiled version of the *Caenorhabditis elegans* wiring diagram, we solve for the optimal layout of 279 nonpharyngeal neurons. In the optimal layout, most neurons are located close to their actual positions, suggesting that wiring minimization is an important factor. Yet some neurons exhibit strong deviations from "optimal" position. We propose that biological factors relating to axonal guidance and command neuron functions contribute to these deviations. We capture these factors by proposing a modified wiring cost function.

Caenorhabditis elegans | optimal placement

diagram and powerful placement algorithms borrowed from computer engineering (29–33). We consider 279 neurons (pharyngeal and unconnected neurons excluded) of the hermaphrodite worm, whose identity, locations of cell bodies, sensory endings, and neuromuscular junctions, as well as the wiring diagram, have been well studied and found to be largely reproducible from animal to animal (34, 35). The length of the worm is >10 times greater than its diameter, allowing us to reduce the problem into one dimension.

By minimizing the cost of connecting the nervous system, our solution predicts the position of most neurons along the anterior–posterior (AP) body axis of the nematode worm. This result suggests that wiring minimization is a good general description of the relationship between connectivity and neuron placement. A comparison of the cost-minimized layout with actual neuron posi-

www.pnas.org/cgi/doi/10.1073/pnas.0506808103 PNAS March 21, 2006 vol. 103 | no. 12 | 4723–4728

- "Plainly, the faculty of language was not instantaneously inserted into a mind/brain with the rest of its architecture fully intact. But we are asking how well it is designed on that counterfactual assumption. How much does the abstraction distort a vastly more complex reality?"

  - N. Chomsky

An *Evo-Devo* Approach

During the last two decades evolutionary developmental biology has become a major research programme whose findings put into question some concepts lying at the core of the Synthetic Theory.

However, some authors are waiting for a 'revolution' in biology, one in which the existing genetic determinism will give way to a new conceptual understanding of the complexity of living organisms.

This interdisciplinary approach is focused on how changes in development bring about evolutionary changes to account for the past and present diversity of morphologies and body plans.

S. Urdy and R. Chirat: Snail shell coiling (re-)evolution and the evo-devo revolution.

Canalization (C. Waddington)

- "… development is robust to changes in genotype and environment."

  M. L. Siegal and A. Bergma. Waddington's canalization revisited: developmental stability and evolution.

- "Individuals within a wild population show remarkably little morphological variation, given the amount of environmental variation they encounter during development and the amount of genetic variation within the population. This phenotypic constancy led to the proposal that individuals were somehow buffered, or canalized, against genetic and environmental variation."

  J. E. Niven. Channelling evolution – canalization and the nervous system.

‘Soft’ Modularity

- “Modularity, a biological approach that views organisms as the integration of partially independent, interacting units at several hierarchical levels, has been described as ‘a conceptual framework for evo-devo’, and ‘a meeting place for evolutionary and developmental biologists’.”

B. K. Hall and W. M. Olson eds.: *Keywords & Concepts in Evolutionary Developmental Biology*.

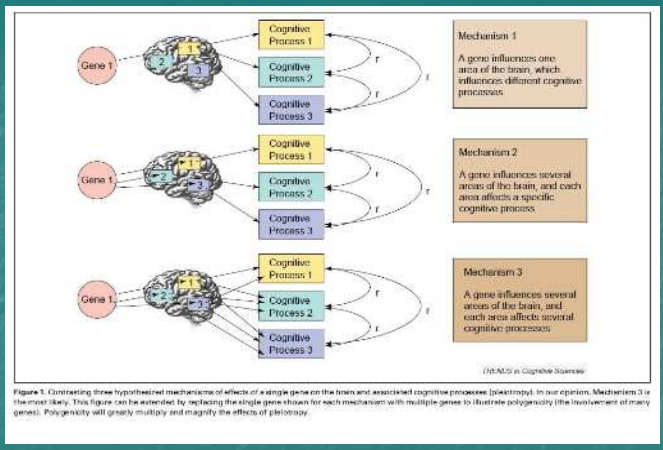


Figure 1. Contrasting three hypothesized mechanisms of effects of a single gene on the brain and associated cognitive processes (pleiotropy). In our opinion, Mechanism 3 is the most likely. This figure can be extended by replacing the single gene shown for each mechanism with multiple genes to illustrate polygenicity (the involvement of many genes). Polygenicity will greatly multiply and magnify the effects of pleiotropy.

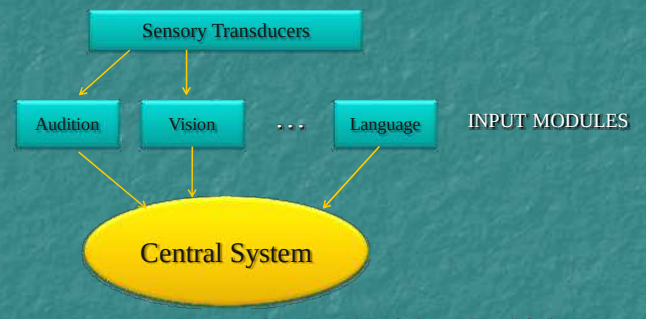
Y. Kovas and R. Plomin 2006. Generalist genes: implications for the cognitive sciences. *Trends in Cognitive Sciences* 10.

Modular Architecture of the Mind

- Domain-Specificity
- Informational Encapsulation
- Autonomous
- Innate
- Mandatory
- Fast
- Deterministic
- Neural Localization
- Idiosyncratic Pathological Breakdown

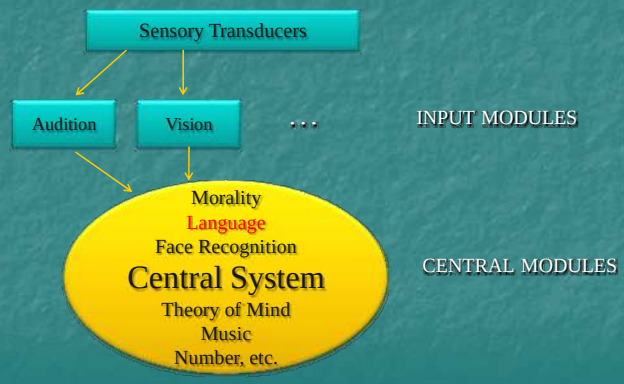
	Central System?	Adaptation?
Fodorian Module	No	No
Chomskyan Module	Yes	No
Darwinian Module	Yes	Yes

Fodorian Modularity



N. Smith 2003. Dissociation and modularity: Reflections on language and mind.

Chomskyan Modularity



Against Strong Innateness

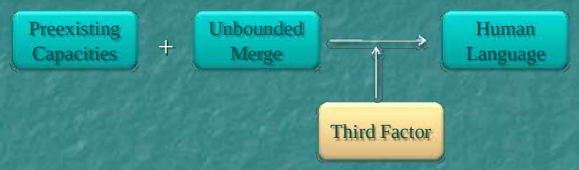
- Departure from strong genetic determinism in *Evo-Devo* and in MP
- "The third factor" in general biological design

Faculty of Language, Broad and Narrow

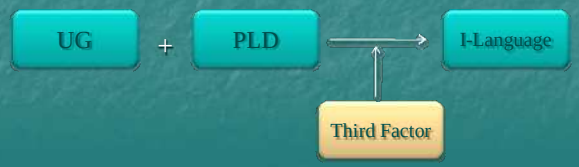
FLN: unique to humans and human language
Recursion only

FLB: not unique to humans and human language
Sensory-Motor and Conceptual-Intentional systems

- Instantaneous Model of Language Evolution



- Instantaneous Model of Language Development



- "… no clear evidence for languages that demonstrably lack recursion of any kind."

  B. Heine and T. Kuteva. 2007. *The Genesis of Grammar*. OUP.

  cf. D. Everett. Cultural constraints on grammar and cognition in Pirahã. Current Anthropology 46. (2005)

(1) student film committee program office

(2) John's friend's friend's friend's friend

- "… unbounded Merge is not only a genetically determined property of language, but also unique to it."

  - N. Chomsky

Serial Verbs:

(1) Me fo kadegbe gba. (Ewe)
I hit lamp break
'I hit-break the lamp.'

(2) Ozó ghá suà àkhé dè. (Edo)
Ozo will push pot fall
'Ozo will push-fall the pot.'

Complex V-V predicates:

(3) John-ga mado-wo tataki-watta. (Japanese)
John-Nom window-Acc hit-broke
'John hit-broke the window.'

Emonds' Paradox

"The concepts *F* of specifically human syntax are precisely those that we might associate with non-human primate cognition."

"The semantic concepts *f* that seem characteristic of humans are not used in human syntax."

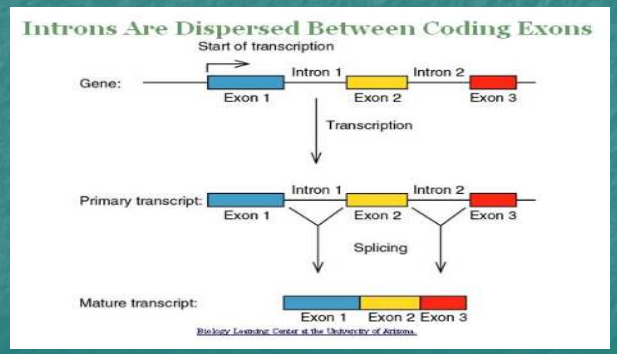
J. Emonds. 2004. What humans have that animals don't have.

The Evolution of [-Interpretable] Features?

- Lexical and Functional Categories
- Exons and Introns

Introns may derive from exons through a process akin to grammaticalization (semantic bleaching).

RNA Splicing



Cell Language and Human Language

"... it may be suggested that human language is ultimately founded in cell language and that human language can be viewed as a transformation of cell language."

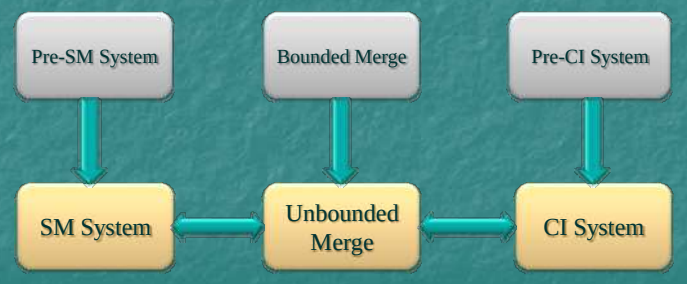
"... a complete understanding of the nature of DNA requires applying the principles of human language to biology."

S. Ji. Isomorphism between cell and human languages: molecular biological, bioinformatic and linguistic implications.

"… to understand better human language, we can also be helped along by a better understanding of the language of the cell."

L. Jenkins. *Biolinguistics.*

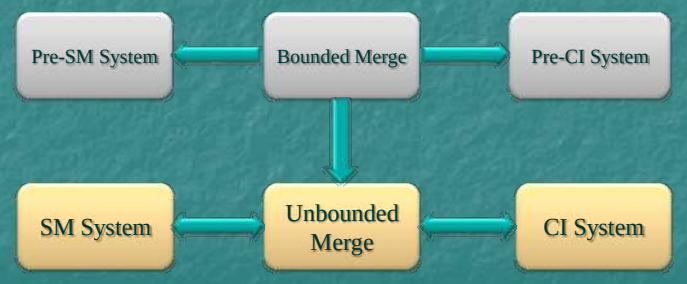
Scenario #1



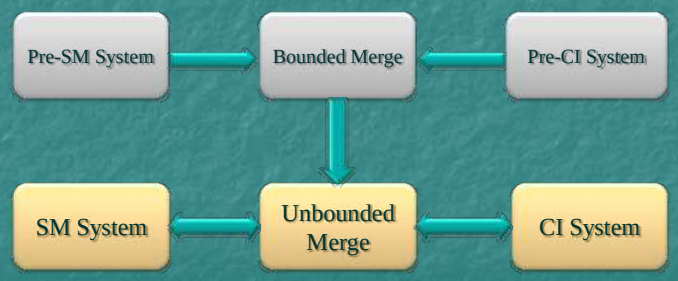
- Language evolution boils down to the emergence of:
  - Unbounded Merge
  - Interfaces
  - Phases (Phase Impenetrability Condition)

  etc.

Scenario #2



Scenario #3



No Precursors to Unbounded Merge?

"… for both evolution and development, there seems to be little reason to suppose that there were precursors to unbounded Merge."

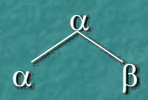
- N. Chomsky

Decomposing Merge

- Merge (α, β) = {α, β}

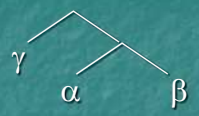
- Embed (α, {α, β}) = {α, {α, β}}

endocentricity ←

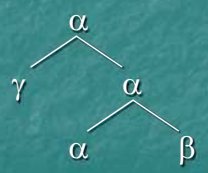


(Fukui 2008)

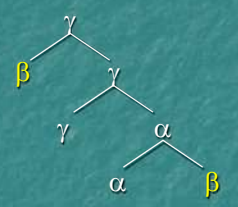
- Recursive Merge (without Embed)



- Recursive Embed



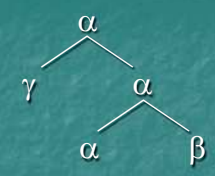
- Internal Merge (Move) + Embed



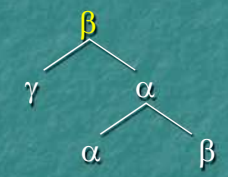
- Why not β for direct Embed without Move?

  Embed (β, {γ, {α,{α, β}}})

- Local Embed



- Non-local Embed



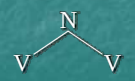
exocentricity ←

Exocentric compounding

(1) Katta-maketa -wa docchi-demo yoi.
won-lost -Top whichever is-good
'Whether we won or lost doesn't matter.'

(2) Tatemono-no takai-hikui-ga juuyoo da.
building-Gen high-low -Nom important is
'The height of the building matters.'

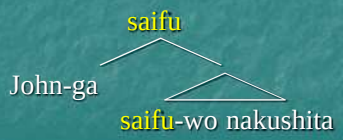
(1)



(2)

- Internally-headed relatives:

(1) [John-ga saifu-wo nakushita no]-wo Mary-ga mitsuketa.
[John-Nom wallet-Acc lost Comp]-Acc Mary-Nom found
'Mary found the wallet John had lost.'

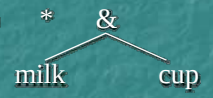


Labeling Two Word Utterances

(1) no label

(2) endocentric

(3) * &



(*in the sense of 'milk & cup')

Possible Precursors to (Bounded) Merge

- Syllable Structure
  - Birdsong
  - Music
- Social Intelligence
  - Theory of Mind (ToM)
  - Machiavellian Intelligence
- Navigation and Foraging
- Number
- Manual Dexterity, Motor Control
  - Action Grammar

Action Grammar

- Pairing Method
- Pot Method
- Subassembly Method

P. M. Greenfield:
Language, tools, and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *BBS* 14 (1991).
Language, tools, and brain revisited. *BBS* 21. (1998)

I. Pairing Method

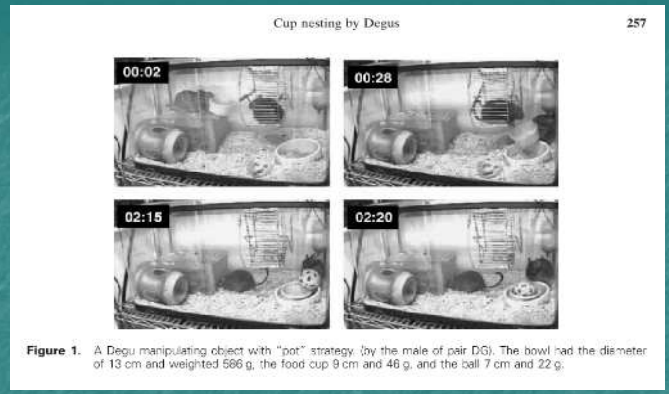
II. Pot Method

- Merge (saw, Mary) = {saw, Mary}
- Merge (John, {saw, Mary}) = {John, {saw, Mary}}

III. Subassembly Method

- Merge (saw, Mary) = {saw, Mary}
- Merge (the, boy) = {the, boy}
- Merge ({the, boy}, {saw, Mary})
  = {{the, boy}, {saw, Mary}}

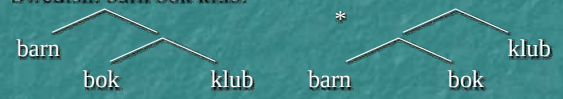
Subassembly Method required



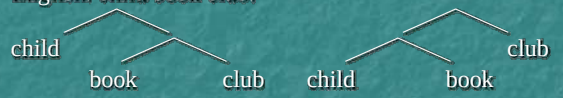
N. Tokimoto and K. Okanoya: Spontaneous construction of "Chinese boxes" by Degus (*Octodon degu*): A rudiment of recursive intelligence? *Japanese Psychological Research* 46 (2004).

Subassembly in Root Compounding

Swedish: barn bok klub:



English: child book club:



T. Roeper and W. Snyder. 2005. Language learnability and the forms of recursion.

Subassembly and Chunking

- Phase = derivational chunk
- Phase Impenetrability Condition:
  - Once formed, chunks cannot be unpacked.

Major Issues

- From Pot to Subassembly?
- From Subassembly to Internal Merge (Move)?
- From bounded to unbounded Merge?

Objections

- Recursion is not limited to humans and human language (such as vision).
- Human language has other components than recursion (such as lexicon).

S. Pinker and R. Jackendoff 2005.

"If future empirical progress demonstrates that FLN represents an empty set, so be it."

Fitch, Hauser and Chomsky 2005.

Lexicon as a Conceptual Barrier

- The existence of a generative lexicon in human language poses a serious challenge to the recursion only hypothesis.
- Does the lexicon belong to FLB or FLN?

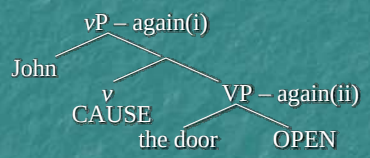
Anti-Lexicalism

- Words are generated by recursive syntax.
- Lexicon decomposed into FLN (recursion) and FLB (SM/CI)
- C-I interface optimized
- There is *virtually* no lexicon.

Syntactic Nature of 'Lexical' Verbs

(1) John opened the door again.
  i. repetitive reading
  ii. restitutive reading

(2)



(3) LCS: [*x* CAUSE [*y* OPEN again(ii)] again(i)]

Ditransitives

(1) a. John gave Mary a book.
b. [$_{vP}$ John *v* [$_{VP}$ Mary V a book]]
c. [*J.* CAUSE [*M.* HAVE *B.*]]

(2) a. John gave a book to Mary.
b. [$_{vP}$ John *v* [$_{VP}$ a book V to Mary]]
c. [*J.* CAUSE [*B.* GO to *M.*]]

- Mapping between syntactic structure and conceptual structure is straightforward.

Evidence from Developmental Data

CAUSE (2;0.4) ≥ HAVE (2;0.7) ≥
Double Obj verbs (2;1.6) >
GO (2;4.0) ≥ Dative Obj verbs (2;4.9)

J. Viau 2006. *Give* = CAUSE + HAVE/GO: Evidence for early semantic decomposition of dative verbs in English child corpora. *BUCLD* 30.

Merge in Early Grammar

- "No verb is an island."
  (cf. Tomasello's *Verb Island Hypothesis*)
- "Children start to use Merge already with their very first word combinations."

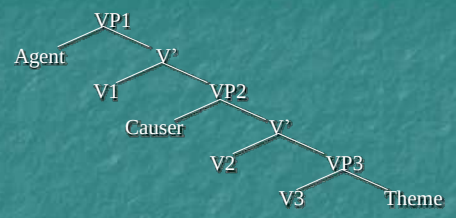
A. Ninio. 2006. *Language and the Learning Curve.* OUP.

(1) *Give*-type verbs have a caused possession interpretation in both variants:
cf. *Where did you give the book?

(2) *Send*-type verbs have a caused motion interpretation in the dative variant:
cf. Where did you send the book?

M. Rappaport Hovav and B. Levin 2008.

Three-Layered Split VP



cf. [x DO [x CAUSE [y BECOME …]]]

K. Fujita:
Middle, ergative and passive in English: A minimalist perspective, *MITWPL* 22. (1994)
Double objects, causatives and derivational economy, *LI* 27. (1996)

(1) This glass breaks easily.
[TP this glass T [μP μ [VP1 IMP V1 [VP2 V2 [VP3 breaks this glass]]]]]

(2) This glass suddenly broke.
[TP this glass T [VP1 V1 [μP μ [VP2 IMP V2 [VP3 breaks this glass]]]]]

Middles	implicit Agent	Generically quantified	+stative
Ergatives	(implicit Causer)	Existentially quantified	+eventive

- *tham/hây* causatives in Thai:

(1) *Saakhaa tham kracok tœœk dooy taŋcay.
Saka cause mirror break by intend

(2) Saakhaa hây dek win dooy taŋcay.
Saka have child run by intend

(3) Saakhaa tham hây kaw?ii lom dooy taŋcay.
Saka cause have chair fall by intend

R. Vichit-Vadakan 1976. The concept of inadvertence in Thai periphrastic causative constructions, in M. Shibatani ed. *Syntax and Semantics* 6: *The Grammar of Causative Constructions*. Academic Press.

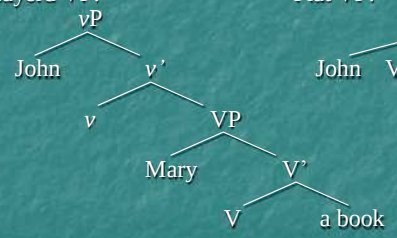
"Causes are realized in a position that is asymmetrically c-commanded by the Agent position."

L. Travis 2005. Agents and Causes in Malagasy and Tagalog, in *The Syntax of Aspect*. OUP.

Simpler Syntax? (Culicover and Jackendoff 2005)

John gave Mary a book.
CS: [x CAUSE [y HAVE z]]

Layerd VP:



Flat VP:

- Flat VP: optimal for SM-system
  - Language for communication
  - Adaptationism
  - Lexicalism
- Layerd VP: optimal for CI-system
  - Language for thought
  - Nonadaptationism
  - Anti-Lexicalism (syntax for thought everywhere)

"It cannot be true literally that 'In the beginning was the word': on the contrary, in the beginning was the sentence."

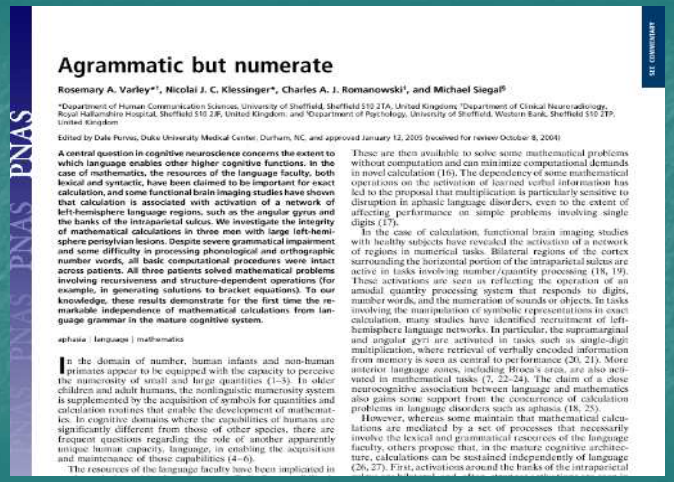
J. Bronowski 1967. Human and animal languages.

- In the beginning was recursion.

Merge to Successor Function

Merge (1,1) = 2
Merge (2,1) = 3, etc.

Mathematical capacity is an abstraction from linguistic operations.



PNAS

SEE COMMENTARY

Agrammatic but numerate

Rosemary A. Varley*†, Nicolai J. C. Klessinger*, Charles A. J. Romanowski‡, and Michael Siegal§

*Department of Human Communication Sciences, University of Sheffield, Sheffield S10 2TA, United Kingdom; †Department of Clinical Neuroradiology, Royal Hallamshire Hospital, Sheffield S10 2JF, United Kingdom; and ‡Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TP, United Kingdom

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved January 12, 2005 (received for review October 8, 2004)

A central question in cognitive neuroscience concerns the extent to which language enables other higher cognitive functions. In the case of mathematics, the resources of the language faculty, both lexical and syntactic, have been claimed to be important for exact calculation, and some functional brain imaging studies have shown that calculation is associated with activation of a network of left-hemisphere language regions, such as the angular gyrus and the banks of the intraparietal sulcus. We investigate the integrity of mathematical calculations in three men with large left-hemisphere perisylvian lesions. Despite severe grammatical impairment and some difficulty in processing phonological and orthographic number words, all basic computational procedures were intact across patients. All three patients solved mathematical problems involving recursiveness and structure-dependent operations (for example, in generating solutions to bracket equations). To our knowledge, these results demonstrate for the first time the remarkable independence of mathematical calculations from language grammar in the mature cognitive system.

aphasia | language | mathematics

In the domain of number, human infants and non-human primates appear to be equipped with the capacity to perceive the numerosity of small and large quantities (1–5). In older children and adult humans, the nonlinguistic numerosity system is supplemented by the acquisition of symbols for quantities and calculation routines that enable the development of mathematics. In cognitive domains where the capabilities of humans are significantly different from those of other species, there are frequent questions regarding the role of another apparently unique human capacity, language, in enabling the acquisition and maintenance of those capabilities (4–6).

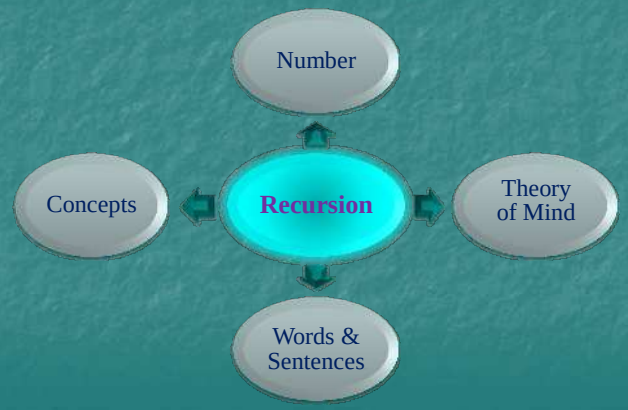
The resources of the language faculty have been implicated in

These are then available to solve some mathematical problems without computation and can minimize computational demands in novel calculation (16). The dependency of some mathematical operations on the activation of learned verbal information has led to the proposal that multiplication is particularly sensitive to disruption in aphasic language disorders, even to the extent of affecting performance on simple problems involving single digits (17).

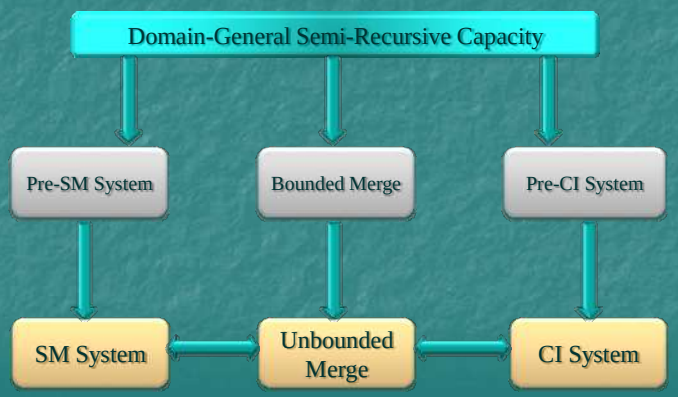
In the case of calculation, functional brain imaging studies with healthy subjects have revealed the activation of a network of regions in numerical tasks. Bilateral regions of the cortex surrounding the horizontal portion of the intraparietal sulcus are active in tasks involving number/quantity processing (18, 19). These activations are seen as reflecting the operation of an amodal quantity processing system that responds to digits, number words, and the numeration of sounds or objects. In tasks involving the manipulation of symbolic representations in exact calculation, many studies have identified recruitment of left-hemisphere language networks. In particular, the supramarginal and angular gyri are activated in tasks such as single-digit multiplication, where retrieval of verbally encoded information from memory is seen as central to performance (20, 21). More anterior language zones, including Broca's area, are also activated in mathematical tasks (7, 22–24). The claim of a close neurocognitive association between language and mathematics also gains some support from the concurrence of calculation problems in language disorders such as aphasia (18, 25).

However, whereas some maintain that mathematical calculations are mediated by a set of processes that necessarily involve the lexical and grammatical resources of the language faculty, others propose that, in the mature cognitive architecture, calculations can be sustained independently of language (26, 27). First, activations around the banks of the intraparietal

Recursion: *The* Generative Engine of the Mind



Scenario #2 (Descent with Modification)



Conclusions (highly tentative)

- MP provides an *Evo-Devo* framework for the study of language evolution.
- Language was adaptive primarily as a cognitive tool, later co-opted for communication.
- Unbounded Merge evolved in several steps, stemming from Action Grammar.
- Human cognitive modules emerged from basic recursive capacity via descent with modification.

Thank you.

Now, on with the Ghost Walk …