Four Ways from Universal to Particular:
How Chomsky’s Language-Acquisition Faculty is Not Selectionist

David Ellerman
Visiting Scholar
Philosophy Department
University of California at Riverside

October 28, 2014

Abstract

Following the development of the selectionist theory of the immune system, there was an attempt to characterize many biological mechanisms as being "selectionist" as juxtaposed to "instructionist." But this broad definition would group Darwinian evolution, the immune system, embryonic development, and Chomsky’s language-acquisition mechanism as all being "selectionist." Yet Chomsky’s mechanism (and embryonic development) are significantly different from the selectionist mechanisms of biological evolution or the immune system. Surprisingly, there is an abstract way using two dual mathematical logics to make the distinction between genuinely selectionist mechanisms and what are better called "generative" mechanisms. This note outlines that distinction.

Contents

1 Introduction 2
2 Selectionist mechanisms 2
3 Selectionist vs. generative approaches to universal grammar 2
4 The two dual forms of mathematical logic 4
5 The two lattices of subsets and partitions 5
6 Four ways to go from universal to particular 7
   6.1 The selectionist mechanism: From universe set to a subset 8
   6.2 The creationist mechanism: From empty set to a subset 9
   6.3 The identification mechanism: From discrete partition to a partition 9
   6.4 The generative mechanism: From indiscrete partition to a partition 9
7 Generative versus selectionist mechanisms 11
1 Introduction

There is a simple, abstract, and logical way to classify four ways that "universals" can give rise to "particulars." One might expect such an abstract classification to have little power to make interesting or significant distinctions. Yet it does give a clear differentiation between what would be called a selectionist mechanism and the generative mechanism of Noam Chomsky's language-acquisition faculty or universal grammar (UG). The purpose of this note is to spell out that distinction.

2 Selectionist mechanisms

There is a long tradition, growing out of biological thought, to juxtapose "selectionist" mechanisms to "instructionist" (or Lamarckian) mechanisms ([28], [22], [11], [9], [10]). Originally the distinction was drawn in rather general terms. In an instructionist or Lamarckian mechanism, the environment would transmit detailed instructions about a certain adaptation to an organism, while in a selectionist mechanism, a diverse variety of (perhaps random) variations would be generated, and then some adaptations would be selected by the environment but without detailed instructions from the environment.

The selectionist-instructionist juxtaposition gained importance with the development of the selectionist theory of the immune system. [21] There is some "generator of diversity" that generates a wide variety of possible adaptations, and then interaction with the environment differentially amplifies some possibilities while the others languish, atrophy, or die off. For instance, in the case of the human immune system, "It is estimated that even in the absence of antigen stimulation a human makes at least $10^{15}$ different antibody molecules--its preimmune antibody repertoire." ([1] p. 1221) quoted in [20, p. 187]) Gerald Edelman has sharpened the selectionist definition and generalized its application.

The long trail from antibodies to conscious brain events has reinforced my conviction that evolution, immunology, embryology, and neurobiology are all sciences of recognition whose mechanisms follow selectional principles. ...All selectional systems follow three principles. There must be a generator of diversity, a polling process across the diverse repertoires that ensue, and a means of differential amplification of the selected variants. [12, p. 7367]

In particular, Edelman develops a selectionist theory of brain development.

[T]he theoretical principle I shall elaborate here is that the origin of categories in higher brain function is somatic selection among huge numbers of variants of neural circuits contained in networks created epigenetically in each individual during its development; this selection results in differential amplification of populations of synapses in the selected variants. In other words, I shall take the view that the brain is a selective system more akin in its workings to evolution than to computation or information processing. [11, p. 25]

The key point is that the possibilities must be in some sense actualized or realized (e.g., as antibodies in low concentration in the immune system) in order for selection to operate on and differentially amplify or select some of the actual variants while the others languish, atrophy, or die off.

3 Selectionist vs. generative approaches to universal grammar

What would the child’s language-learning faculty look like if it was selectionist in this sense? There is a naïve and a sophisticated selectionist account of language acquisition. In the naïve account,
the child would (perhaps randomly) generate a diverse range of babblings, some of which would be differentially reinforced by the linguistic environment [36].

Skinner, for example, was very explicit about it. He pointed out, and he was right, that the logic of radical behaviorism was about the same as the logic of a pure form of selectionism that no serious biologist could pay attention to, but which is [a form of] popular biology – selection takes any path. And parts of it get put in behaviorist terms: the right paths get reinforced and extended, and so on. It’s like a sixth grade version of the theory of evolution. It can’t possibly be right. But he was correct in pointing out that the logic of behaviorism is like that [of naïve adaptationism], as did Quine.[7, p. 53]

As noted, Willard Van Orman Quine adopted essentially this approach to language learning.

An oddity of our garrulous species is the babbling period of late infancy. This random vocal behavior affords parents continual opportunities for reinforcing such chance utterances as they see fit; and so the rudiments of speech are handed down. [35, p. 73]

It remains clear in any event that the child’s early learning of a verbal response depends on society’s reinforcement of the response in association with the stimulations that merit the response, from society’s point of view, and society’s discouragement of it otherwise. [35, p. 75]

Since language users can generate a variety of rule-based grammatical sentences never before spoken, the child would have to rather miraculously generalize the grammatical rules from the reinforced variants. In order to be adequate as an explanation, a model needs to take seriously the speaker’s rule-based competency.

There is a sophisticated version of a selectionist model for the language-acquisition faculty or universal grammar (UG) which could be called the format-selection (FS) approach (Chomsky, private communication). The diverse variants that are actualized in the mental mechanism are different sets of rules or grammars. Then given some linguistic input from the linguistic environment, the grammars are evaluated according to some evaluation metric, and the best rules are selected.

Universal grammar, in turn, contains a rule system that generates a set (or a search space) of grammars, \( \{G_1, G_2, \ldots, G_n\} \). These grammars can be constructed by the language learner as potential candidates for the grammar that needs to be learned. The learner cannot end up with a grammar that is not part of this search space. In this sense, UG contains the possibility to learn all human languages (and many more). ... The learner has a mechanism to evaluate input sentences and to choose one of the candidate grammars that are contained in his search space. [31, p. 292]

The idea is that after a sufficient stream of linguistic inputs, the mechanism would converge to the best grammar that matches the linguistic environment. Since it is optimizing over sets of rules, this model at least takes seriously the need to account for rule-based competency. Early work (through the 1970s) on accounting for the language-acquisition faculty or universal grammar (UG) seems to have assumed such an approach.

The earliest ideas were roughly as follows. Suppose that UG provides a certain format for languages, that is, a specification of permitted types of rules and permissible interactions among them. Any rule system satisfying the proposed format qualifies as a possible human language. ... The mind employs certain primitive operations to interpret some of the data presented to it as linguistic experience, then selects among the languages consistent with this experience in accordance with an evaluation metric that assigns an abstract value to each language. [4, p. 52]
The problems that eventually arose with the FS approach could be seen as the conflict between descriptive and explanatory adequacy. In order to describe the enormous range of human language grammars, the range of grammars considered would make for an unfeasible computational load of evaluating the linguistic experience. If the range was restricted to make computation more feasible, then it would not explain the variety of human languages.

It was an intuitively obvious way to conceive of acquisition at the time for – among other things – it did appear to yield answers and was at least more computationally tractable than what was offered in structural linguistics, where the alternatives found in structural linguistics could not even explain how that child managed to get anything like a morpheme out of data. But the space of choices remained far too large; the approach was theoretically implementable, but completely unfeasible. [7, p. 173]

Instead of the format-selection (FS) approach, the alternative principles and parameters (P&P) approach ([3], [6], [5]) to universal grammar was then developed:

we no longer consider UG as providing a format for rule systems and an evaluation metric. Rather, UG consists of various subsystems of principles; it has the modular structure that we regularly discover in investigation of cognitive systems. Many of these principles are associated with parameters that must be fixed by experience. The parameters must have the property that they can be fixed by quite simple evidence, because this is what is available to the child; the value of the head parameter, for example, can be determined from such sentences as John saw Bill (versus John Bill saw). Once the values of the parameters are set, the whole system is operative. [4, p. 146]

Our purpose here is to give an abstract conceptual differentiation of the P&P approach from the sophisticated selectionist approach of the FS system (not to mention from the crude selectionism in behaviorism or naïve Darwinism).

Jerry Fodor and Massimo Piattelli-Palmarini took a different approach to that differentiation. In the process, they generated some controversy with evolutionary biologists by claiming that "Skinner’s account of learning and Darwin’s account of evolution are identical in all but name” or, to be more precise, ”what is wrong with Darwin’s account of the evolution of phenotypes is very closely analogous to what is wrong with Skinner’s account of the acquisition of learned behavior.” [17, p. xvi] They emphasize aspects of what is broadly called the ”neo-neo-Darwinism,” Evo Devo, or the ”extended evolutionary synthesis” [34]. Instead of wading into that controversy, we take the different approach in this note of showing how selectionist mechanisms (e.g., FS) and ”generative” P&P-type mechanisms can be differentiated at a very abstract logico-mathematical level. Hence we must turn to a recent development in mathematical logic.

4 The two dual forms of mathematical logic

George Boole [2] originally developed what might be called Boolean logic as the logic of subsets, not the logic of propositions. The interpretation solely in terms of propositions and the name propositional logic came later.

The algebra of logic has its beginning in 1847, in the publications of Boole and De Morgan. This concerned itself at first with an algebra or calculus of classes, to which a similar algebra of relations was later added. Though it was foreshadowed in Boole’s treatment of ”Secondary Propositions,” a true propositional calculus perhaps first appeared from this point of view in the work of Hugh MacColl, beginning in 1877. [8] pp. 155-156]
When Boolean logic is interpreted as the logic of subsets, then variables stand for subsets of some given universe set $U$, the operations are subset operations, and a (subset-)valid formula or (subset-)tautology is a formula so that no matter what subsets of $U$ are substituted for the variables, the whole formula will evaluate to $U$ for any non-empty $U$. It is then a theorem (known to Boole), not a definition, that it suffices to consider the case where $U = 1$ is a singleton which has only two subsets 1 and $\emptyset$ (the empty set). Hence validity in the special case of propositions with the two truth-values 1 and 0, i.e., truth-table validity, is equivalent to general subset-validity. Eventually, the special case of propositional variables came to dominate so truth-table validity became the definition of a tautology rather than a theorem about subset-validity (see any contemporary logic textbook).

What is lost by this focus on the special case of propositional logic rather than the general case of subset logic? Around the middle of the twentieth century, the theory of categories was formalized [13] and an older informal notion of duality in algebra was formalized as the reverse-the-arrows duality of category theory [26]. The older informal duality in algebra was the juxtaposition of subgroups to quotient groups, subrings to quotient rings, and in general subobjects to quotient objects—which in the basic case of sets was the juxtaposition of subsets to quotient sets (the latter being equivalent to equivalence relations or partitions on a set). For instance, F. William Lawvere calls the general notion of a subobject a "part" and "The dual notion (obtained by reversing the arrows) of 'part' is the notion of partition." [25, p. 85] Hence when the special case of "propositional" logic is seen as the general logic of subsets, then the idea arises of there being a dual logic of quotient sets or partitions ([15], [16]). That idea of a dual logic does not arise when subset logic is seen only as propositional logic since "propositions" do not have a category-theoretic dual.

## 5 The two lattices of subsets and partitions

The two logics of subsets and partitions are represented algebraically by the Boolean algebra of subsets of a universe $U$ and the algebra of partitions on a universe set $U$ ($|U| \geq 2$). For our purposes, it suffices to consider the two lattices, the familiar Boolean lattice of subsets of $U$ (where the partial order is inclusion) and the lattice of partitions on $U$ where the partial order is the "refinement" relation between partitions.

A partition $\pi = \{B\}$ on a universe $U$ is a set of nonempty blocks $B$ that are disjoint and whose union is $U$. Given two partitions $\pi = \{B\}$ and $\sigma = \{C\}$ on $U$, the partition $\pi$ refines the partition $\sigma$, written $\sigma \preceq \pi$, if for every block $B \in \pi$, there is a block $C \in \sigma$ such that $B \subseteq C$. Figure 1 illustrates the two lattices for the universe $U = \{a, b, c\}$ (where the partial order is indicated by the lines).

![Figure 1: Lattices of subsets and partitions](image)

Each lattice has a top and bottom and that gives us the four universals in our title. In the Boolean lattice of subsets, the top is the universe set $U$ and the bottom is the null set $\emptyset$. In the lattice of partitions, the top is the discrete partition $1 = \{\{u\} : u \in U\}$ where all blocks are singletons, and the bottom is the indiscrete partition $0 = \{\{U\}\}$ with only one block consisting of the universe $U$. 

5
The duality between subsets of a set and partitions on a set extends to the extensive analogies between the elements of a subset and the distinctions of a partition, where a distinction or dit of a partition \( \pi = \{ B \} \) on \( U \) is an ordered pair \((u, u') \in U \times U\) of elements in different blocks of \( \pi \). There is a set-theoretic representation of the lattice of partitions where each partition is represented by its set of distinctions or ditset:

\[
\text{dit} (\pi) = \{(u, u') : \exists B, B' \in \pi; B \neq B'; u \in B; u' \in B'\},
\]

and where the partial order is just inclusion between ditsets since:

\[
\sigma \preceq \pi \iff \text{dit} (\sigma) \subseteq \text{dit} (\pi).
\]

The complement of a ditset is the set of indistinctions of the partition:

\[
\text{indit} (\pi) = U \times U - \text{dit} (\pi) = \{(u, u') : \exists B \in \pi; u, u' \in B\}
\]

which is simply the equivalence relation associated with the partition. The ditsets of partitions on \( U \) are thus the complements of equivalence relations on \( U \) and they might be called the partition relations on \( U \).

Given any subset \( S \subseteq U \times U \), its reflexive-symmetric-transitive or rst closure \( \text{cl} (S) \) is the smallest equivalence relation containing \( S \) (which is well-defined since the intersection of two equivalence relations is an equivalence relation). But it might be noted that this closure operation is not a topological closure operation since the union of two rst-closed sets is not necessarily rst-closed. The interior \( \text{int} (S) \) of a subset \( S \subseteq U \times U \) is the complement of the rst-closure of the complement, i.e., \( \text{int} (S) = cl (S^c)^c \), so it is the ditset of some partition. To define the partition operation corresponding to any logical subset operation (e.g., union, intersection, conditional, etc.), apply the subset operation to the ditsets of the partitions, take the interior of the result, and then the partition corresponding to that interior. For instance, the meet \( \pi \land \sigma \) of two partitions \( \pi \) and \( \sigma \) may be defined by the ditset:

\[
\text{dit} (\pi \land \sigma) = \text{int} [\text{dit} (\pi) \cap \text{dit} (\sigma)].
\]

Thus we can take any formula of subset logic and interpret it as a formula of partition logic. The atomic variables would represent partitions on \( U \) instead of subsets of \( U \). Given such an interpretation of a formula \( \Phi (\pi, \sigma, \ldots) \), an member \( u \in U \) being an element of the subset represented by \( \Phi (\pi, \sigma, \ldots) \) is analogous to an ordered pair \((u, u') \) being a distinction of the partition represented by \( \Phi (\pi, \sigma, \ldots) \). The two definitions of a valid formula are also analogous. A formula \( \Phi (\pi, \sigma, \ldots) \) is a valid formula of subset logic, i.e., a tautology, if for any subsets of \( U \) substituted for the variables, the formula evaluates to the set of all possible elements \( U \) (the top of the lattice) for any \( U \) \(|U| \geq 1\). Similarly, a formula \( \Phi (\pi, \sigma, \ldots) \) is a valid formula of partition logic, i.e., a partition tautology, if for any partitions on \( U \) substituted for the variables, the formula evaluates to the partition that makes all possible distinctions, i.e., the top-of-the-lattice discrete partition \( 1 \) with the ditset \( \text{dit} (1) = U \times U - \Delta \) (\( \Delta \) is the diagonal \( \{(u, u) : u \in U\}\)), for any \( U \) \(|U| \geq 2\).

The following Figure 2 summarizes the dual relationships between the two logics (see [15] or [16] for more on partition logic).
Four ways to go from universal to particular

In the two lattices of the dual logics, there are four universals, the tops and bottoms of the two lattices. The four logico-mathematical ways to abstractly characterize going from universal to particular are the four ways of going from one of the universals to a particular subset or partition in the lattice. In view of the duality between elements and distinctions, the four ways can be characterized as: (1) killing off elements, (2) creating elements, (3) killing off distinctions, and (4) creating distinctions.

If \( S \) represents any particular subset of \( U \) and \( \pi \) represents any particular partition on \( U \), then the four ways are:

1. the selectionist mechanism \( U \to S \): to go from the universe set \( U \) to a particular subset \( S \) by "selecting" the elements of \( S \) and eliminating or "killing off" the elements of the complement \( S^c \);
2. the creationist mechanism \( \emptyset \to S \): to go from the empty set \( \emptyset \) to a particular subset \( S \) by "creating" the elements of \( S \);
3. the identification (or classification or symmetry-making) mechanism \( 1 \to \pi \): to go from the discrete partition \( 1 \) to a particular partition \( \pi \) by identifying elements ("killing off" distinctions) of \( U \) (in a consistent way); and
4. the generative (or symmetry-breaking) mechanism \( 0 \to \pi \): to go from the indiscrete partition \( 0 \) to a particular partition \( \pi \) by "generating" distinctions on \( U \) (in a consistent way).

The four schemes can be related in terms of duals (the elements-distinctions duality) and opposites (all versus none) as in Figure 3.

---

**Figure 2: Dual analogies between the subset and partition logics**

| Subset Logic | Partition Logic |
|--------------|----------------|
| 'Elements'   | Distinctions (u,u') of a partition \( \pi \) |
| All 'elements' | Universe set \( U \) |
| No 'elements' | Empty set \( \emptyset \) |
| Duality      | Indiscrete partition \( \emptyset \) (no dits) |
| Subsets are images \( f() \) of injections \( f:S\to U \) | Partitions are inverse-images \( f^{-1}() \) of surjections \( f:U\to T \) |
| Formula variables | Subsets of \( U \) |
| Logical operations | Partitions on \( U \) |
| \( \cup, \cap, \Rightarrow, \ldots \) | Partition ops. = Interior of subset ops. applied to dissets |
| Formula \( \Phi(\pi,\sigma,\ldots) \) holds at 'element' | Element \( u \) is in subset \( \Phi(\pi,\sigma,\ldots) \) |
| \( \Phi(\pi,\sigma,\ldots) \) | Pair \((u,u')\) is a distinction of partition \( \Phi(\pi,\sigma,\ldots) \) |
| Valid formula | Valid formula |
| \( \Phi(\pi,\sigma,\ldots) = U \) for any subsets \( \pi,\sigma,\ldots \) of any \( U \) (\( |U| \geq 1 \)) | \( \Phi(\pi,\sigma,\ldots) = 1 \) for any partitions \( \pi,\sigma,\ldots \) on any \( U \) (\( |U| \geq 2 \)) |

---

6 Four ways to go from universal to particular
A selectionist \((U \rightarrow S)\) mechanism and a generative \((0 \rightarrow \pi)\) mechanism are related by taking the dual and the opposite (in either order).

### 6.1 The selectionist mechanism: From universe set to a subset

The first scheme \(U \rightarrow S\) is the abstract logico-mathematical model of a selectionist process since it starts with an actualized set of diverse alternatives \(U\) and then a number of the alternatives are eliminated by some fitness criterion or evaluation metric while the remaining alternatives are selected (e.g., by differential amplification).

The original example of a selectionist process is Darwinian evolution where the set of diverse alternatives is generated over time by random genetic mutations and then the environment applies a fitness filter (see [17] on the general generate-and-filter idea). In the application to the immune system, the actualized set of diverse alternatives are the generated set of antibodies in low concentration and then the selectionist process differentially amplifies those that fit an invading antigen. Gerald Edelman’s various selectionist models [11] also fit this scheme. One of Peter Medawar’s metaphors for a selectionist scheme was a jukebox where all the tunes were already actualized as records in the jukebox and then one was selected. In the format-selection or FS approach to universal grammar, the mental mechanism must generate some representation of the diverse variety of grammars, and then a chunk of linguistic experience is evaluated according to some evaluation metric to find the best fit among the various systems of rules.
6.2 The creationist mechanism: From empty set to a subset

The second scheme \( \emptyset \rightarrow S \) is the abstract logico-mathematical model of a creation story where elements are, in effect, created out of nothing. While this is one type of "creation story," the Big Bang creation theory [32] is modelled not by this \( \emptyset \rightarrow S \) scheme but by the generative \( 0 \rightarrow \pi \) scheme where the making of distinctions is rendered as symmetry-breaking. The \( \emptyset \rightarrow S \) scheme is perhaps the least interesting to model actual processes.

![Subset lattice](image)

Figure 5: From the empty set \( \emptyset \) to a particular subset \( S \)

6.3 The identification mechanism: From discrete partition to a partition

The third scheme \( 1 \rightarrow \pi \) is the abstract logico-mathematical model of any classification, partitioning [24, p. 82], or quotienting process that proceeds by making consistent identifications ("consistent" means that the identifications must be reflexive, symmetric, and transitive to form an equivalence relation).

![Partition lattice](image)

Figure 6: From the discrete partition \( 1 \) to a particular partition \( \pi \)

For instance, the classification of animals to species where \( 1 \) represents each animal by itself and \( \pi \) represents the partition of the set of animals as to species. Mathematically, the action of a group on a set is automatically reflexive, symmetric, and transitive so it defines an equivalence relation where the equivalence classes are called "orbits" [27, p. 99]. This \( 1 \rightarrow \pi \) scheme is "symmetry-making" while the opposite scheme \( 0 \rightarrow \pi \) is "symmetry-breaking."

6.4 The generative mechanism: From indiscrete partition to a partition

The fourth scheme \( 0 \rightarrow \pi \) is the abstract logico-mathematical model of any *generative* process where a number of different outcomes (represented by the blocks of \( \pi \)) can be generated by consistently
making distinctions ("consistent" means nothing can be distinguished from itself, distinguishing must be symmetric, and if \( u \) is distinguished from \( u' \) and \( u = u_1, u_2, ..., u_n = u' \), then one of the pairs \( (u_i, u_{i+1}) \) must also be distinguished for \( i = 1, ..., n-1 \), all of which means the set of distinctions must be anti-reflexive, symmetric, and anti-transitive, i.e., a partition relation).

\[
\begin{align*}
\{\{a\},\{b\},\{c\}\} \\
\{\{a\},\{b\}\},\{\{a\},\{c\}\} \rightarrow \{\{b\},\{c\}\} \\
\{\{a\},\{b\}\},\{\{a\},\{c\}\},\{\{b\},\{c\}\} \\
\{\{a,b,c\}\}
\end{align*}
\]

**Partition lattice**

Figure 7: From the indiscrete partition 0 to a particular partition \( \pi \)

The most "atomic" type of distinction is a binary partition of the single block \( \{U\} \in 0 \) into two blocks and then the binary partitions can be combined or joined together. The join \( \pi \lor \sigma \) of two partitions \( \pi = \{B\} \) and \( \sigma = \{C\} \) is the partition of nonempty intersections \( B \cap C \) (or, in terms of ditsets, \( \text{dit}(\pi \lor \sigma) = \text{dit}(\pi) \cup \text{dit}(\sigma) \)), where the interior is not needed since a union of partitions relations is always a partition relation). And the most "efficient" binary partition is one that divides the block \( \{U\} \) into two equal parts (assuming an even number of elements). The classic example is where \( U \) has \( 2^n \) elements which can be enumerated using \( n \)-place binary numbers. Then \( U \) can be divided into two equal parts by the binary partition according to whether the \( i^{th} \) binary digit is 0 or 1. The join of those binary partitions for \( i = 1, ..., n \) would go from the indiscrete partition 0 all the way to the discrete partition 1, so the \( n \) equal-binary partitions are Shannon’s \( n \) bits [14].

Often a \( 0 \rightarrow \pi \) generative process proceeds not only by joining binary partitions (with not necessarily equal blocks) but by designating one of the blocks as in the game of twenty questions where the block with the yes-answer to the yes-or-no question is designated. In this case, the \( 0 \rightarrow \pi \) process goes not just from the indiscrete partition 0 to a particular partition \( \pi \) but from the single block \( \{U\} \in 0 \) to a specific block \( B \in \pi \) (like a correct answer in the game of twenty questions) by following the yes-branches on the binary tree. This is the case of the generative mechanism that is of most interest for our purposes.

A binary partition with a designated block is just a choice with two options, and it might be represented by a switch with a neutral setting (representing the state before the choice is made) and then two options such as a Left Option and a Right Option.

**Figure 8: Switch to go from Neutral to either Left or Right Options**
For instance consider the example where $U$ has $8 = 2^3$ elements represented by the 3-digit binary numbers $b_3b_2b_1$. There would be three switches for $i = 1, 2, 3$ where the two options for the $i^{th}$ switch are a 0 or 1 in the $i^{th}$ place $b_i$ of the 3-digit binary number. Then the initial state is the indiscrete partition $0 = \{ U \}$ where all the switches are in neutral. The first switch (controlling the first digit on the right) determines the binary partition with one block having the four elements $b_3b_20$ (Left Option) and the other block having the four elements $b_3b_21$ (Right Option), and so forth for the other two switches. When all three switches have been set one way or the other, that determines the transition from the single block $\{ U \} \in 0$ to a specific singleton block represented by a specific 3-digit binary number.

Our principal application of the $0 \rightarrow \pi$ generative scheme is Chomsky’s principles and parameters (P&P) description of the language-acquisition faculty or universal grammar ([3], [6], [5]). In our simple model, the parameters are represented by the switches that can be moved to the left or right (from the original setting of neutral) by the child’s linguistic experience, and the underlying principles are expressed in the whole setup defining the grammatical meaning of the left and right settings.

A simple image may help to convey how such a theory might work. Imagine that a grammar is selected (apart from the meanings of individual words) by setting a small number of switches—0, say—either ”On” or ”Off.” Linguistic information available to the child determines how these switches are to be set. In that case, a huge number of different grammars (here, 2 to the twentieth power) will be prelinguistically available, although a small amount of experience may suffice to fix one. [18, p. 154]

Needless to say, this imagery implicitly allows for a neutral setting on the switches (sometimes called the ”initial state $S_0$”) since otherwise the original or ”factory” setting of the switches would determine a specific grammar independent of experience.

Borrowing an image suggested by James Higginbotham, we may think of UG as an intricately structured system, but one that is only partially ”wired up.” The system is associated with a finite set of switches, each of which has a finite number of positions (perhaps two). Experience is required to set the switches. When they are set, the system functions. The transition from the initial state $S_0$ to the steady state $S_s$ is a matter of setting the switches. [4, p. 146]

Another implication of this general type of $0 \rightarrow \pi$ model for Chomsky’s language-acquisition faculty is the interpretation of the adjective ”universal” in the phrase ”universal grammar.” It does not mean a specific grammatical rule common to all languages (which would be a subset-logic interpretation of ”universal”).

The switch-settings of the metaphor above are in Chomsky’s terminology the ”parameters” defined by universal grammar. Notice that this image underscores the sense in which universal grammar, the initial state of the language-learner, need not comprise an account of what languages have in common—to continue the metaphor, different switch-settings could give rise to very different grammatical system. [18, p. 154]

### 7 Generative versus selectionist mechanisms

It might also be useful to illustrate a selectionist and a generative mechanism to solve the same problem of determining one among the $8 = 2^3$ options considered in the last section. The eight possible outcomes might be represented as:

\[
|000\rangle, |100\rangle, |010\rangle, |110\rangle, |001\rangle, |101\rangle, |011\rangle, |111\rangle.
\]

11
In the selectionist scheme, all eight variants are in some sense actualized or realized in the initial state \( S_0 \) so that a fitness criterion or evaluation metric (as in the FS scheme) can operate on them. Some variants do better and some worse as indicated by the type size in Figure 9.

\[
\begin{align*}
|000\rangle, & |100\rangle, |010\rangle, |110\rangle, |001\rangle, |101\rangle, |011\rangle, |111\rangle \\
|000\rangle, & |100\rangle, |010\rangle, |110\rangle, |001\rangle, |101\rangle, |011\rangle, |111\rangle \\
|000\rangle, & |100\rangle, |010\rangle, |110\rangle, |001\rangle, |101\rangle, |011\rangle, |111\rangle \\
|010\rangle, & |010\rangle, |010\rangle, |010\rangle, |010\rangle, |010\rangle, |010\rangle, |010\rangle
\end{align*}
\]

Figure 9: An abstract model of a selectionist learning mechanism

Eventually the "unfit" options dwindle, atrophy, or die off leaving the most fit option \( |010\rangle \) as the final steady state \( S_s \).

In the generative learning scheme, the initial state \( S_0 \) is where all the switches are in neutral so all the eight potential outcomes are in a "superposition" state indicated by the plus signs in the following Figure 10.

**Initial State: All switches at Neutral**

\[
|000\rangle + |100\rangle + |010\rangle + |110\rangle + |001\rangle + |101\rangle + |011\rangle + |111\rangle
\]

- **Left Option at first position**
- **Right Option at second position**
- **Left Option at third position**

Figure 10: An abstract model of a generative learning mechanism

We assume that the initial experience sets the first switch (or first parameter in the P&P model) to the left option which reduces the state to \( |000\rangle + |100\rangle + |010\rangle + |110\rangle \) (where the plus signs between these options indicate that the second and third switches are still in neutral). Then subsequent experience sets the second switch to the right option and the third switch to the left option. Thus we reach the same outcome \( |010\rangle \) as the final state \( S_s \) in the two models but by quite different mechanisms.

We have now differentiated a selectionist \((U \rightarrow S)\) mechanism from a generative \((0 \rightarrow \pi)\) mechanism at a very abstract logico-mathematical level.

There is a whole literature where "selectionist" is interpreted very broadly as non-instructionist so that the generative \((0 \rightarrow \pi)\) mechanism is also described in those overly-broad terms as being "selectionist" (e.g., by describing the generative mechanism as "selecting" switch-settings).
I suggest that some important lessons for linguistics and cognitive science can, indeed, be drawn from contemporary biology, but that the new principles and the new assumptions came to bury learning by instruction and to replace it with learning by selection, a radically different process. What now replaces learning everywhere in biology has nothing to do with a transfer of structure and everything to do with mechanisms of internal selection and filtering affecting a pre-programmed chain of multiple internal recombinations and internal "switches." [33, p. 3]

Now we see that this sort of setting of internal switches is better described as a generative ($0 \rightarrow \pi$) mechanism whereas many of the other "learning" mechanisms in biology (e.g., in the immune system) are correctly described as a selectionist ($U \rightarrow S$) mechanism since the latter involves the actualization of some "universal" repertoire of possibilities some of which are selected.

A hierarchy of genetic switches as in a stem cell or in embryonic development would be a generative mechanism ([30], [19]). One might imagine a hypothetical selectionist mechanism to replace stem cells that would postulate low concentrations of the different types of cells through the body, so that, say, muscle cells would be selected to multiply in a muscle environment while the other types of cells would be inactive there. Yet what is found biologically is not that type of selectionist mechanism but the generative mechanism of stem cells (where the muscle environment sets the switches to produce a muscle cell—in addition to reproducing the stem cell).

Peter Metawar [28] explains the selectionist-instructionist juxtaposition by contrasting a jukebox (with the musical records taken as internal) with a record player (with the records taken as external). The jukebox has a set of pre-existing options one of which is selected by the simple pushing of a button whereas when a record player plays music, the set of external instructions must be supplied in the form of a record. Thus a jukebox is a genuine selectionist ($U \rightarrow S$) mechanism. Medawar also describes the development of the embryo as being selectionist:

Embryonic development... must therefore be an unfolding of pre-existing capabilities, an acting-out of genetically encoded instructions; the inductive stimulus is the agent that selects or activates one set of instructions rather than another. [29, p. 295]

But in terms of our differentiation, embryonic development is a generative ($0 \rightarrow \pi$), not a selectionist ($U \rightarrow S$), mechanism. In a similar manner, it is easy to see that Gerald Edelman’s [11] various models of brain development and learning are all selectionist ($U \rightarrow S$) mechanisms.

In this manner, one could go over all the examples broadly called "selectionist" and see which were genuinely selectionist ($U \rightarrow S$) mechanisms and which were generative ($0 \rightarrow \pi$) mechanisms—which shows the surprising fruitfulness of the quite abstract logico-mathematical differentiation between $U \rightarrow S$ and $0 \rightarrow \pi$ mechanisms.

References

[1] Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson. 1994. Molecular Biology of the Cell. New York: Garland.

[2] Boole, George 1854. An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities. Cambridge: Macmillan and Co.

[3] Chomsky, Noam. 1981. Lectures on Government and Binding. Dordrecht: Foris Publications.

[4] Chomsky, Noam. 1986. Knowledge of Language: Its Origin, Nature, and Use. New York: Praeger.

[5] Chomsky, Noam. 1995. The Minimalist Program. Cambridge MA: MIT Press.
[6] Chomsky, Noam, and Howard Lasnik. 1993. The Theory of Principles and Parameters. In Syntax: An International Handbook of Contemporary Research, edited by J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann, 506–69. Berlin: de Gruyter. (reprinted in: Chomsky 1995)

[7] Chomsky, Noam, and James McGilvray, eds. 2012. The Science of Language: Interviews with James McGilvray. Cambridge UK: Cambridge University Press.

[8] Church, Alonzo. 1956. Introduction to Mathematical Logic. Princeton NJ: Princeton University Press.

[9] Cziiko, Gary. 1995. Without Miracles: Universal Selection Theory and the Second Darwinian Revolution. Cambridge: MIT Press (A Bradford Book).

[10] Dennett, Daniel. 1995. Darwin’s Dangerous Idea: Evolution and the Meanings of Life. New York: Touchstone.

[11] Edelman, Gerald M. 1987. Neural Darwinism: The theory of neuronal group selection. Basic Books, New York.

[12] Edelman, Gerald M. 2004. Biochemistry and the Sciences of Recognition. Journal of Biological Chemistry. 279 (9): 7361–69.

[13] Eilenberg, S. and S. Mac Lane. 1945. General Theory of Natural Equivalences. Transactions of the American Mathematical Society. 58, No2, 231-94.

[14] Ellerman, David 2009. Counting Distinctions: On the Conceptual Foundations of Shannon’s Information Theory. Synthese. 168 (1 May): 119-149.

[15] Ellerman, David 2010. The Logic of Partitions: Introduction to the Dual of the Logic of Subsets. Review of Symbolic Logic. 3 (2 June): 287-350.

[16] Ellerman, David 2014. An Introduction of Partition Logic. Logic Journal of the IGPL. 22, no. 1: 94–125.

[17] Fodor, Jerry, and Massimo Piattelli-Palmarini. 2010. What Darwin Got Wrong. New York: Farrar, Straus and Giroux.

[18] Higginbotham, James. 1982. Noam Chomsky’s Linguistic Theory. Social Research 49 (1): 143–57.

[19] Jacob, François. 1973. The Logic of Life. Translated by Betty Spillman. New York: Pantheon.

[20] Jenkins, Lyle. 2000. Biolinguistics: Exploring the Biology of Language. New York: Cambridge University Press.

[21] Jerne, Niels. 1955. The natural selection theory of antibody formation. Proc. National Academy of Sciences U.S.A. 41, 849.

[22] Jerne, Niels 1967. Antibodies and learning: Selection versus instruction. In The neurosciences: A study program. G. C. Quarton, T. Melnechuk and F. O. Schmitt eds., New York: Rockefeller University Press: 200-5.

[23] Jerne, Niels 1985. The Generative Grammar of the Immune System. Science. 229 (4718): 1057–59.

[24] Lawvere, F. William and Stephen Schanuel 1997. Conceptual Mathematics: A first introduction to categories. New York: Cambridge University Press.
[25] Lawvere, F. William and Robert Rosebrugh 2003. *Sets for Mathematics*. Cambridge MA: Cambridge University Press.

[26] Mac Lane, Saunders 1971. *Categories for the Working Mathematician*. Springer Verlag; New York.

[27] Mac Lane, Saunders, and Garrett Birkhoff. 1967. *Algebra*. New York: Macmillan.

[28] Medawar, Peter B. 1960. *The Future of Man: Reith Lectures 1959*. London: Methuen.

[29] Medawar, Peter B. 1982. *Pluto’s Republic*. Oxford UK: Oxford University Press.

[30] Monad, Jacques 1974. *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*. New York: Collins/Fontana.

[31] Nowak, M. A., and N. L. Komarova. 2001. Towards an Evolutionary Theory of Language. *Trends in Cognitive Sciences*. 5 (7 July): 288–95.

[32] Pagels, Heinz. 1985. *Perfect Symmetry: The Search for the Beginning of Time*. New York: Simon and Schuster.

[33] Piattelli-Palmarini, Massimo 1989. Evolution, Selection and Cognition: From ‘Learning’ to Parameter Setting in Biology and in the Study of Language. *Cognition*. 31 (1 Feb.): 1–44.

[34] Pigliucci, Massimo, and Gerd B. Müller, eds. 2010. *Evolution: The Extended Synthesis*. Cambridge MA: MIT Press.

[35] Quine, Willard Van Orman. 2013. *Word & Object (New Edition)*. Cambridge MA: MIT Press.

[36] Skinner, B. F. 1976. *About Behaviorism*. New York: Vintage Books.