As soon as we talk about syntactic structures that enable semantic interpretations, we have crossed over from the domain of *signs* into that of *symbols* as combinatorial syntactic elements that primarily map to one another rather than to an item or occurrence in the environment/world. Understanding semantics-enabling syntactic symbols in this way does not risk a hypostatization of abstract entities, since they are still constrained and embodied in pattern-governed regularities or pictures. But the manner of this constraining or incarnating regime is not a matter of the simple compositionality of pictures. Pictures do not map to syntactic symbols, but only to the interrelations between symbols which, in themselves, have no mapping relation whatsoever to pattern-regularities in the world. In other words, pattern-governed regularities in the real order are caught up in the relations between symbols, not the other way around. Such relations themselves are not pattern-governed in the sense of belonging to the real order. They are instead already rule-governed (i.e., combinatorial) relations belonging to a different order, the autonomous order of symbols. Collapsing the distinction between signs and symbols, picture-mappings and symbolic syntactic interrelations, or regarding conceptual activities as 'a species of pattern-governed behavior'[234] is a recipe for all sorts of confusions, one to which Sellars himself has unfortunately contributed some ingredients.

Just because conceptual activities and their undergirding symbolic interrelations are built on pictures and iconic-indexical signs, this does not mean that conceptual activities and symbols are *made of* pattern-governed regularities and mapping relations. Being built on something is not the same as being made of something. Let us clarify these points in the context of our thought experiment.

234 'The distinction between pattern-governed behavior and rule-governed activity is not a difference in kind; rather, rule-governed activity is a species of pattern-governed behavior: a recursive loop generated through the interaction between complex patterns.' R. Brassier, 'Transcendental Logic and True Representings', *Glass Bead* 0 (2016), <http://www.glass-bead.org/article/transcendental-logic-and-true-representings/>.

SIGN-VEHICLES AND SYMBOL-DESIGN

In the multi-agent system, the automata have their own inductive moves. The contact of the fuzzy grey item with the heap of black (E_i–E_j) changes their wiring diagram. The state of the representational system or the wiring diagram of the automata is now a second-order isomorphism—i.e., an equivalence (but not equality) relation—that registers as E_i*–E_j*, a statistical inductive association with its corresponding transitions and obstructions. For example, the occurrence of the fuzzy item making contact with the heap of black and then a shrieking noise is registered on $\mathbb{K}$'s wiring diagram as a pattern-governed regularity that permits certain rudimentary inductive moves: Whenever N_s (shrieking noise) at $p + \Delta p$ *and* $t + \Delta t$ then G (fuzzy grey item) at $p + \Delta p$ and $t + \Delta t$. It should be pointed out that registering the noise N_s depends on a few factors. Firstly, it must be of some interest or significance for $\mathbb{K}$, i.e., it must play a role for $\mathbb{K}$'s overall behavioural economy. If the noise is of no interest to $\mathbb{K}$, it will not register on $\mathbb{K}$'s 'attentional system' (Dehaene). Only the top-down attentional amplification of modular processes can mobilize the change in the wiring diagram and make it available to $\mathbb{K}$'s global workspace or awareness of the environment. Secondly, registering the noise N_s crucially depends on the sufficiency of $\mathbb{K}$'s causal structure or wiring for pattern recognition. Without this criterion, $\mathbb{K}$ would not be able to differentiate the amplitude and frequency of the shrieking noise from the rustling noise and would therefore be incapable of making inductive moves—transitions and obstructions—related to these items.

Our assumption, however, is that $\mathbb{K}$, $\mathbb{S}$, and $\mathbb{M}$ satisfy such criteria. They are equipped with the adequate causal structure to differentiate not only N_s and N_r, but also such noises of interest to them insofar as they play an important role in their behavioural (or perception-action) ecology. Upon occurrence of the pattern N_s, $\mathbb{K}$'s auditory system registers it as an input acoustic pattern σ which is associated with the occurrence of N_s and, correspondingly, with the fuzzy item touching the black item. The automaton tags this acoustic pattern σ by emitting a whistling sound or acoustic cue Σ using its inbuilt electromechanical devices. The acoustic cue Σ has no meaning in the sense of describing N_s or signifying what E_i*–E_j* (the picture of the fuzzy item touching the black item accompanied by a shrieking noise) is. From now on, whenever $\mathbb{K}$ registers σ it emits the whistling sound Σ. Within the multi-agent system, $\mathbb{K}$'s repeated reuse of Σ is also registered by $\mathbb{S}$ and $\mathbb{M}$ as a cue for the occurrence of N_s and, correspondingly, the pattern-governed regularity $E_i^*–E_j^*$.

The significance of the relation between $E_i–E_j$, $E_i^*–E_j^*$, N_s and σ in triggering Σ is simply that of the statistical frequency of spatiotemporal occurrences registered by the structurally sufficient and behaviourally active automata. In short, it is devoid of complex inferential or conceptual relationships. This is precisely what Peirce calls an indexical sign, a sign whose interpretant is entirely a statistical regularity of some occurrence (a causal or temporal invariance) and whose interpreter requires only a heuristic device to interpret it. An index is a sign that is somehow causally linked with something else in space and time. Furthermore, indexical sign-vehicles such as Σ are built on iconic signs that concern the resemblance of one pattern to another. This resemblance, however, is wholly a matter of an *arbitrary* interpretation of the similarity-vagueness of something versus something else. In this sense, iconicity is really the stimuli-based discrimination of *stuff*.

Pictures qua mapping functions are indexical signs, which themselves are built on icons or stimuli-based discriminations. In terms of their reference, both indices and icons are arbitrary and generic. A clarifying example of indexical sign-vehicles would be social animal alarm calls. In sighting a predator, a member of the group vocalizes a specific sound that is statistically associated with the presence of predators. This sound or indexical sign does not, however, specify the type of the predator, nor does it relay any information to other members of the group regarding the predator's exact location. It simply signals the presence of a danger in their vicinity. Similarly, the only significance of Σ is that it communicates an *indexical* one-to-one relationship between the interpreter (the agent) and the interpreted ($E_i–E_j$). Transmitting and receiving Σ is what we can call 'communication', reserving the term for this rudimentary schema of transmitting and receiving an *indexical* one-to-one relationship between a sound cue and an interpreted event or series of events.

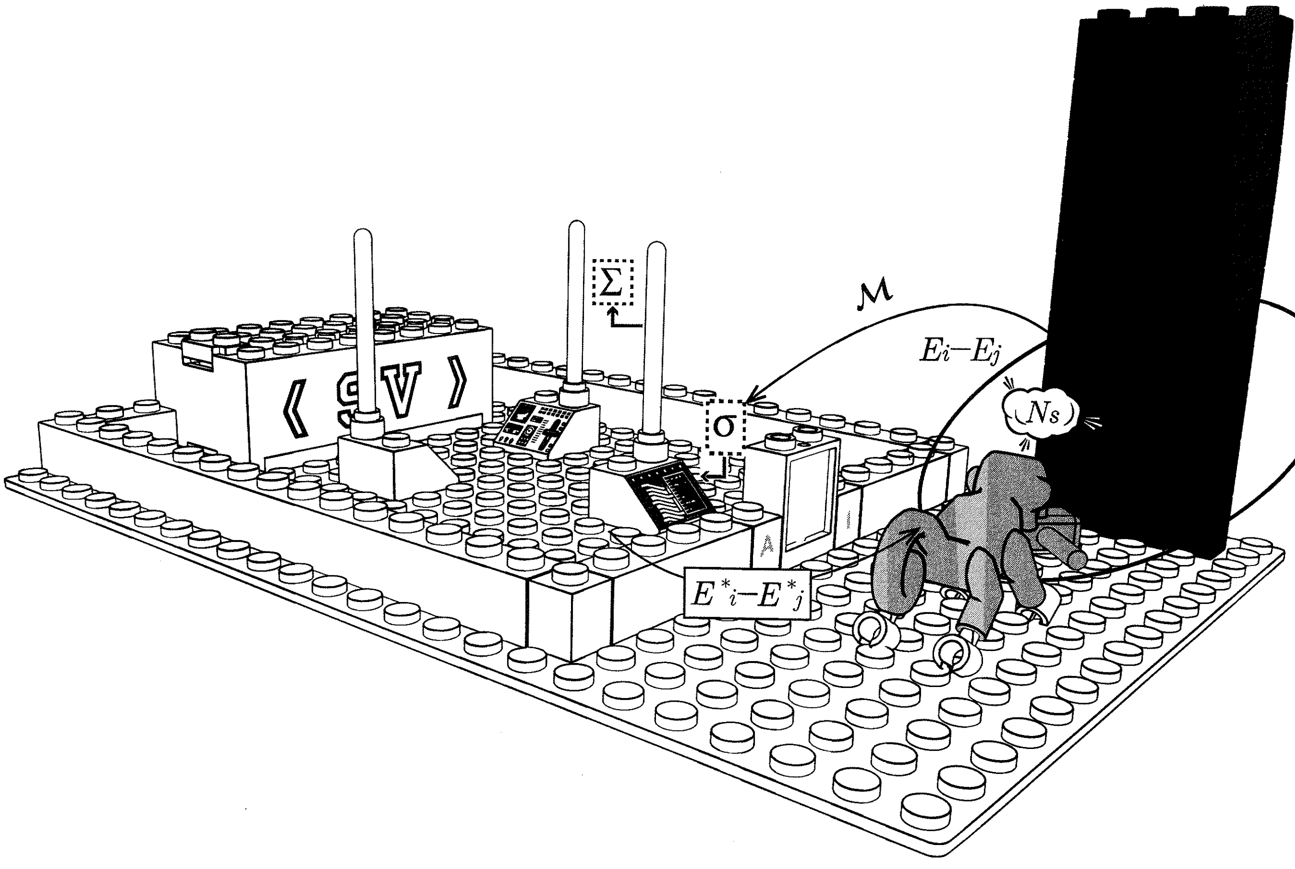
Upon registering an input acoustic pattern, $\mathbb{K}$'s internal pattern-recognition mechanisms match it with $\mathbb{K}$'s archive of *input patterns* correlated with previously catalogued patterns of occurrences. If the registered acoustic pattern matches σ, the automaton produces Σ as a signal pertaining to the alleged occurrence of the associated events E_i–E_j. Setting aside the problem of the arbitrariness of the representational correspondence between σ and the E_i–E_j pattern of occurrence, the indices σ for $\mathbb{K}$ and Σ for $\mathbb{S}$ and $\mathbb{M}$ are fundamentally narrow in the range of their relationships with other indices associated with other pattern-governed regularities.

The insufficiency and unreliability of the prevalent indexical sign-vehicles is nowhere more manifest than in natural phenomena involving mimicry and deception. Take for instance the parasitoid blue butterfly *phengaris rebeli*, known for its intricate social parasitism on a species of ants called *myrmica* ants. The butterfly lays eggs close to the ant colony. Its broods discharge the same chemical signals by which ants distinguish their own. By mimicking the ants' indexical chemical sign, the butterfly broods trick the ants into carrying them into their nest. Once they are in the nest and as they mature, by mimicking the acoustic signals the queen ant uses to mobilize worker ants to bring food to it, the butterfly larvae climb the social hierarchy of the ant colony. Associating the mimicked acoustic signal with their queen, the worker ants begin to bring food to the parasitoid larvae, in the process starving their queen—the only ant that can detect the larvae as aggressors—to death. However, for the *myrmica* ants the semiotic nightmare does not end here, for the oversaturation of chemical signals as the result of the activity of both the ants and *p. rebeli*'s parasitism lures yet another parasitoid to the ant colony—the *ichneumon* wasp cited by Darwin as evidence that shakes the very idea of a 'beneficent and omnipotent God'.[235]

Owing to their one-to-one mapping structure, just like indices, pictures qua nonconceptual representings are limited in terms of how they relate to or interact with other pictures. The complex relations between pictures or

---

235 C. Darwin, *The Life and Letters of Charles Darwin* (2 vols. New York: Appleton, 1898), vol.2, 105.



Automata and nonconceptual representational mapping: The sound N_s is mapped onto the automaton's devices as an acoustic pattern σ. N_s is statistically correlated with the pattern-governed regularity E_i–E_j. As the automaton registers σ, it produces the sound Σ to tag the occurrence of the pattern E_i–E_j. The representational mapping $\mathcal{M}$ (E_i*–E_j* to E_i–E_j) is not a meaning relation. It is a transformation between two natural objects, the environmental state and the internal state of the structurally sufficient and behaviourally active automaton. For this reason, the sound cue does not in any semantic sense *signify* an aspect of the external environment, nor does it describe any detail about the internal state of the automaton itself. It is simply a probabilistic marker for the statistical convergence of the automaton's σ-state and the pattern of occurrence for external events E_i–E_j.

pattern-governed regularities cannot be obtained by pictures themselves. Claiming that automata can, *in principle*, fully structure and navigate their environment simply by adequate picturing inevitably results in *picturing regress*. This is because the structuration and navigation of the environment requires the selection, explicitation, and elaboration of diverse relations between pattern-governed regularities. In order for such complex relations to be singled out, a new picture or one-to-one mapping must be acquired. But in so far as the acquisition of a new picture or the detection of a new pattern-governed regularity results in additional relations between pictures and patterns which the new picture cannot by itself index, still further pictures must be acquired. This process can go on ad infinitum without ever covering the complex relations or interactions between patterns, hence the regress. Where there are one-to-many and many-to-many correlations between patterns, the picturing capacity, as a one-to-one mapping, always falls short.

Moreover, from the practical perspective, even if the automata could indeed fully structure and navigate their environment by adequate and complete picturing, from a computational point of view this would be unfeasible, since for every matter-of-factual property of items or patterns of occurrences of events in the world, they would have to acquire not only a picture or representational mapping but also additional pictures to represent the diverse relations between those pictures and their corresponding regularities. With this ever-expanding repertoire of picturings, the size of the automata's internal model would grow. Since computational resources and memory would need to be allocated to these pictures in order to support the structure of the internal model, as it increased in size the internal model would become less and less effectively computable. Once the complexity of the internal model's structure could no longer be increased owing to the lack of effective computability and resource starvation, it would no longer be capable of singling out or picturing new phenomena and items. The representational competencies of the automata would then begin to diminish. This is a ubiquitous scenario among any species equipped only with iconic and indexical sign-vehicles, and not symbols.

For indexical sign-vehicles to cover the diverse relations between pattern-governed regularities, they have to be incorporated into a *different kind* of semiotic system—one whose signs are not in a one-to-one or representational relation to referents, patterns, or events. The order of picturing and indexical signs is necessary but not sufficient to cover the interactions between patterns. This sufficiency can only be obtained by a different kind of semiotic system that is not composed of mappings. This is a system in which the semiotic elements do not represent or relate to a pattern-governed regularity, an external event, or a substantive referent, but only to one another. The combinatorial order of symbols that permits syntactic configurations is exactly this inter-related semiotic system.

In the vernacular use of the concept of symbol, symbols are symbols because they are inscriptions, letters, characters, sounds, or gestures that abstractly represent something else. But such elements are not symbols in the strict syntactic and formal sense. Symbols are only symbols in so far as they are *abstract* nonrepresentational entities that relate or refer to one another rather than to something external to them (i.e., a represented target event, pattern, etc). For semiotic elements to be reorganized as symbols certain criteria must be satisfied, such as the discretization of signs and the establishing of a combinatorial structure (i.e., rules for the possibility and impossibility of combining signs) whereby recursive processes can generate new and more complex relations between abstract elements.

It is the order of symbols that has syntactic structure, not the order of signs—whether iconic or indexical. And it is only the syntactic structure of symbols—elements that only stand in one-to-many and many-to-many relations to one another—that can codify relationships between pattern-governed regularities, pictures, or indexical signs. Conceptual semantic activities are *sui generis* pattern -governed regularities (or rules) that pattern or structure pattern-governed regularities—i.e., single out and elaborate the diverse relationships between patterns—because they are undergirded by syntactic structures that structure and codify such relationships. The codifying syntactic structures are themselves afforded by symbols, semiotic elements that are defined not by how they represent an item or an external relation of reference, but by the combinatorial relationships in which they stand in relation to one another.

Put simply, the order of symbols is not a representational order in the sense that iconic or indexical signs are, but an order whose systematic inter-relations can codify or structure *indexical relationships*— something that indexical signs by themselves cannot accomplish. Indices are accordingly caught up in the combinatorial and codifying relationships between symbols or elements of syntax. Just as indices are built on icons, symbols are built on indices. But, as Terrence Deacon points out, the fact that symbols are built on indices and indices are built on icons does not mean that the relations that hold between them are relations of simple compositionality.[236] Indices are not made of icons and symbols are not made of indices. Symbols, indices, and icons are *distinct* levels of structuring. Whereas representational icons and indices have a limited capacity for structuring, symbols, by virtue of their combinatorial self-referentiality have an in-principle unrestricted structuring capacity. For our automata, the capacity for the full-blown structuration of their world does not begin with their ability to correctly picture their environment, but only with the advent of symbolic technologies that permit the syntactic codification and semantic elaboration of the relations between pictures.

This is, however, not to suggest that the automata may not have more adequate picturing abilities than ours. They may well be furnished with more sophisticated mechanisms for pattern-recognition and algorithms capable of more optimal compression of data and hence more fine-grained picturing—or regularity detection—abilities. This is the fundamental insight of Ray Solomonoff—as formulated in the context of algorithmic complexity and the computational account of inductive inference—that there exists an intrinsic duality of compression and regularity: anything that can compress data is a type of regularity, and any regularity can compress data.[237] For example we might imagine our automaton's picturing abilities to be modelled on

236 T. Deacon, 'Beyond the Symbolic Species', in *The Symbolic Species Evolved* (Dordrecht: Springer, 2012), 13.

237 R. Solomonoff, 'A Formal Theory of Inductive Inference part 1', *Information and Control* 7.1 (1964), 1–22.

better predictive compression algorithms such as Solomonoff's induction or Crutchfield's ε-machine reconstruction, which employ a formal computational (rather than ordinary epistemic) version of Occam's razor and the principle of multiple explanations to yield inductive models smaller in size and of a lower degree of arbitrariness.[238]

Briefly, Solomonoff's predictive induction employs the formal computational definition of simplicity: All knowledge available in a domain at a specific time can be written as a binary string. In this case, a new observation or experience at time t_2 increases the length of the string. Solomonoff's induction problem is then how to predict the next bit of the string by extrapolating the known length of the string. Independently of Andrey Kolmogorov and Gregory Chaitin, Solomonoff showed that there exists an optimal prior probability distribution on any potentially infinite string such that one can compute the next best possible or formally simplest extrapolation using this universal prior distribution. The best predictive model can then be defined in terms of the length of the shortest program that outputs the simplest possible string.[239]

The picturing abilities of our automata could indeed be modelled on such programs.[240] But prediction is not the same as explanation. Even in

238 See J. Crutchfield, 'The Calculi of Emergence: Computation, Dynamics, and Induction', *Physica D*, vol. 75 (1994), 11–54.

239 For more nontechnical details on Solomonoff's theory of formal induction see S. Legg, 'Is There an Elegant Universal Theory of Prediction?', in *Algorithmic Learning Theory* (Dordrecht: Springer, 2006), 274–87, and N. Chater and P. Vitányi, 'Simplicity: A Unifying Principle in Cognitive Science?', *Trends in Cognitive Sciences* 7:1 (2003), 19–22.

240 Crutchfield's ε-machine reconstruction follows much the same principle as Solomonoff's induction except for one significant difference: it takes a string of bits and reverse-engineers the physical states of the system or machine responsible for the generation of that pattern or regularity. Here, the word 'machine' stands for a physical system capable of S possible physical states. The dynamic of such a system can then be defined in terms of the transition between its states. This transition can be thought as an oscillation between state s_1 and state s_2, and hence can be modelled on a program which, at every discrete time stamp, generates bitstrings by

information theory, prediction and explanation are subtly distinguished from one another. Furthermore, as argued earlier with regard to the problems of complete picturing, an agent modelled purely on such formal learning machines would still run into the problem of effective computability. For example, Solomonoff's universal learning machine presupposes an infinite time limit and runs on a universal Turing machine. For a realistic agent with a finite time and limited computational resources, not only may the implementation of Solomonoff's algorithm result in resource impoverishment since the universal Turing machine is the most resource-consuming class of abstract machines, it may also lead to the uncomputability of the halting problem—that is, the undecidability of whether the machine should accept or reject an input to yield an output.

Accordingly, what is required for explanation is not better prediction or the compression-regularity duality, but the ability to selectively compress data or to single out one regularity over another. And precisely what language—starting with the order of symbols—affords agents is the ability to selectively compress data, to single out and elaborate diverse relations between regularities—not merely to picture pattern-governed regularities but to describe and explain them in context. As discussed in the previous chapter, this copula of description and explanation is what material inferences ultimately are. For our automata to count as belonging to the order of general intelligence, they must be able to perform material inferences, to have the practical competences or know-how to use concepts. In short, they must have an artificial general language to syntactically encode the relationships between pictures, to structure such relationships by semantically elaborating their material incompatibility and consequence relations, and ultimately to make

---

emitting 0 or 1 depending upon which state the system is currently in. ε-machine reconstruction takes such a bitstring as its input—that is, it begins with the current state of a finite state machine and predicts the future dynamics of the system based on the history of its past transitions and current state, i.e., it extrapolates whatever information is needed to predict the next bit in the string.

explicit such relations in formal inferences. Coming into possession of such structuring syntactic-semantic abilities, however, entails something more than just the addition of new sign-vehicles or better pictures of the world. It requires symbol-design—the construction of symbolic tokens that can be combined in a variety of ways to form codifying structures and assume semantic roles in the process of being exchanged between agents.

In the following sections, we shall see how the sign-using automata can be endowed with structuring syntactic-semantic abilities by coming into possession of symbols whose combinatorial generative capacities permit the stabilization of inter-agent interactions and the encoding of diverse relationships between pattern-governed regularities. In developing the ability for symbol design, no additional modification to the multi-agent system will be introduced. What we want is to see how the interacting automata themselves, using the resources they already have at their disposal—in particular, the capacity to receive and transmit acoustic signals or sound cues—can acquire symbols.

AN ACOUSTIC EXPERIMENT IN THE PRODUCTION OF SYMBOLS

In chapter 5, a new constraint was incorporated into the modifications of the automata: the automata can only produce *quasi-continuous* sounds. The motivation for the addition of this constraint was that it would allow us to monitor the processes required for the generation of symbols in the toy universe of our automata.

Recall that the automata are provided with organs or electromagnetic devices capable of constricting the continuous sound (noise) and discontinuing it in a specific manner. In this way, the sound is discontinued either abruptly or gradually (i.e., within a temporal window). The constricted or discontinued sound now represents an acoustic range. It is regulated by different constricting devices working in coordination with one another, paralleling the role of the lungs, nasal canal, tongue, and lips in human vocalization. At this point, the sound is still continuous but is also regulated by the manner of discontinuation or constriction.

The quasi-continuous sound can be refined further into a coarsely discretized sound via the frequent reuse (i.e., production and recognition) of the sound as a form of statistical modulation over time. Once our hypothetical agents solve the problem of the discretization of sound cues into stable tokens in interactions whose role can be tracked, diversified, and manipulated, they can single out and relay a wide range of behaviours pertaining to themselves and to the environment. Solving this problem requires the transition of the sound cue (sign) from a partially stabilized and regulated acoustic range (a quasi-continuous sound) to a coarsely discretized sound that can be not only effectively reproduced and recognized by the agents, but also used as a building block for composite symbolic sounds.

In the end what we want to do is to equip the automata with a medium through which they can not only share their positional-perspectival awareness but also, more importantly, compare their perspectives in a stabilized and structured manner, and ultimately arrive at an aperspectival (objective) view of the world. This development would require a stable and combinatorial medium capable of replacing the parochial referential relations between sound cues/signs and occurrences/references (i.e., the representational relation between Σ and $E_i - E_j$) with syntactic and ultimately semantic relations between symbolic sound-tokens. But in order to get on the path of such a development, first the sound cue Σ needs to transform from a *sign* referentially correlated to something external into a *symbol* that primarily stands in relation with other symbols. However, to recapitulate, what we are after in the toy universe of automata is a system comprising a finite repository of abstract elements which *individually* do not represent anything yet which, precisely because of this abstractness, can be combined to create composite elements which then can be put together in conformity with combinatorial rules, thus generating increasingly complex syntactic structures capable of encoding the relations between nonconceptual representing or pattern-governed regularities. The abstract elements of this system are symbols, and to reach the stage where automata would be able to employ symbol-designs to syntactically and semantically structure their world, their acoustic medium of communication will have to undergo a drastic transformation. As we shall see, this transformation is impossible without a process of discretization, which is required in order to implement effective combinatorial mechanisms among sound-tokens.

In the previous sections, we saw that $\mathbb{K}$, $\mathbb{S}$, and $\mathbb{M}$ have the capacity to communicate—in the rudimentary sense of receiving and transmitting sound cues Σ marking pictured pattern-governed regularities. In this communicative regime, however, the instability and the excess of noise in the quasi-continuous sound Σ negatively affects the precision of its signalling function, both in the sense of what it signals (the E_i–E_j pattern of occurrence as opposed to a different occurrence signalled by another quasi-continuous acoustic cue) and how the cue is successfully recognized and consumed in order to result in coordinated group response. An example of the inefficacy of quasi-continuous sound cues for transmission and reception as reliable signals would be animal alarm calls where, because of the instability and fuzziness of the acoustic signal, the call can play different roles (to signal the presence of a predator, as a mating call, in the presence of food, etc.). The appropriate group action is therefore always compromised by the fuzziness of the quasi-continuous sound cue exchanged between the communicator and the communicant. In issuing a certain sound cue, there is no guarantee that the group will *reliably* receive the communicated signal as a cue for the referent targeted by the communicator.

In addition, the quasi-continuous sound possesses a low degree of combinatoriality. Since quasi-continuous acoustic cues are structurally unstable and fuzzy in terms of their acoustic range, multiple-cue integration is severely limited. This lack of combinatoriality means that stable and coded relations between different sound cues cannot be established properly, if at all. Without these relations between sound cues afforded by combinatorial grouping, the role of the acoustic signal in connection to its source (the correspondence between σ and E_i–E_j) and its target (recognition and consumption of the sound cue by other agents) can be neither specified nor stabilized. In isolation, the role of a sound cue as a probability marker is not optimally reliable, either in relation to what it signals or in relation to its potential signal-consumers.

In our toy model, suppose that the solution for reducing the noise and fuzziness of the quasi-continuous sound is the systematic reuse of acoustic cues in the context of interaction between agents. Over time, the reuse of sound cues among the automata in an interactive framework leads to a crystallized distribution of preferred models of sound production between agents in a fashion similar to a self-organizing map or an artificial neural network utilizing unsupervised learning algorithms to cluster data by projecting high-dimensional input data (training samples) onto regular and discretized or low-dimensional data. The decrease in the number and range of acoustic data means that the agents over time automatically select and reuse a small number of clusters of preferred low-dimensional acoustic data. To put it less technically, a model of self-organization can be conceived in the multi-agent system whereby the mere vocalizations of the automata (i.e., the reuse of sound cues or acoustic signals) can converge on a small finite repertoire of preferable coarsely discretized sounds which do not communicate or represent anything, in that they are no longer sign-vehicles but abstract acoustic elements that can be combined into composite sound-tokens.

Structural discretization, together with the decrease in the number of sound clusters, results in the construction of the first building blocks of speech—that is, a finite repertoire of sounds as symbolic tokens.[241] The automata are now in the possession of a limited but well-structured set of reusable and sharable sounds—beeps, clicks, chirps, rustles, etc.—that can be combined in accordance with compositional constraints to form more composite nonrepresentational sounds. These sounds are comparable to the basic phonological units of our speech, and can be combined in conformity with statistical-acoustic transition rules to produce composite sound-tokens that can be exchanged in interactions between automata. For example, in our phonological system the transition from a [V] sound

241 For more details on the design and simulation of models for the discretization of sound into symbolic sound-tokens in an artificial multi-agent system see C. Browman and L. Goldstein, 'Competing Constraints on Intergestural Coordination and Self-organization of Phonological Structures', *Bulletin de la communication parlée* 5 (2000), 25–34, and P.-Y. Oudeyer, *Self-Organization in the Evolution of Speech* (Oxford: Oxford University Press, 2006).

to a [SH] sound is a high-cost transition, whereas [V] to [AA] is a low-cost and optimal transition. The low-cost transition results in a preferable concatenation [VAA] that can be added to the sound repository for reuse and further composition with other sound units. In this way, numerous utterable concatenations can be produced from a very limited set of basic discrete sound units.[242] The discrete units can take on new features within the concatenated compositions. These new features can then be singled out and introduced as additional constraints on the composition of sounds, producing ever more varied and richer concatenations that would constitute the skeleton of the automata's syntactic utterances.

The organization of a sparse number of discretized sounds should not be interpreted as a reduction in the capacity of sounds in their role in the interaction between the automata because, in the new configuration, sounds can be combinatorially grouped and structured. In other words, the process of discretization and the advent of combinatorial capacities go hand in hand with the production of a finite repertoire of sounds which no longer have a communicative signalling function, since they are properly speaking compositional elements of symbolic tokens defined by their syntactic configurations.

The stabilization of acoustic data has a number of important ramifications. It supplies the automata with a repository of sounds that can be retrieved and reproduced. But more fundamentally, it transforms fuzzy

---

242 The difference between basic sound units or elementary phonological objects and the variety of features or properties they can take in sound concatenations is usually defined in terms of the difference between the so-called *emic* and *etic* units. In this case, emic units are phonemes (that which is sounded), while etic units are the observed variant forms of emic units in a system of composite sounds. As etic units, sounds are treated as phonetic elements (i.e., the variations of that which is sounded, or phonemes). The distinction between emic and etic units is crucial for defining sound-tokens qua symbols in terms of invariant and variant forms, i.e., a limited set of stable invariant abstract objects out of which numerous varied abstract features can be constructed. The emic-etic relationships are not exclusive to sounds. They can be extended to every possible symbolic unit including graphemes, morphemes, lexemes, and grammemes.

sound cues into acoustic tokens that can be used as a protocurrency for a new type of interaction wherein the communicative function of sign-vehicles is replaced by symbolic tokens upon which structuring syntactic and semantic abilities can be built. With the introduction into our multi-agent system of symbols in the form of discretized and combinatorial sounds, the automata enter into the syntactic domain of language where the combinatorial relations between symbols can be manipulated to generate more complex syntactic structures capable of encoding broader ranges of relationships between pattern-governed regularities.

GENERATIVE PROCESSES AND THE HIERARCHY OF SYNTACTIC COMPLEXITY

What makes the discretization of sound a consequential structural change in our multi-agent system is that it enables the evolution of speech and cognition for the automata. It simultaneously fulfils the necessary condition for the emergence of speech and transforms the interaction between agents into an abstract system capable of generating increasingly adequate structuring abilities. Such a system can be called language, a geistig scaffold on the basis of which mind can become the unifying point or configuring factor for the structuration of the world, thoughts, and actions.

Owing to their structural stability, which facilitates the sharing and reuse of acoustic tokens, the coarsely discretized sound-composites qua symbols can assume specific and abstract roles in the interaction between the automata. By losing their signalling or communicative function, the discretized sounds become building blocks of symbol-design, i.e., tokens which, depending on how they are exchanged between the automata, can assume different roles ranging from syntactic utterances (structure encoding expressions) to semantic vocabularies (structuring expressions). Having come into possession of sound-tokens or symbols which are defined not by what they stand for but by the abstract roles they play in the interactions between the automata, the multi-agent system can transform from a communicative regime into a linguistic system through which automata can increasingly structure their interaction with their environment by structuring their own interactions.

The transformation of the rudimentary sound cues into a set of limited discrete sound units from which compound sounds can be constructed endows the automaton with combinatorial capacities—specifically, the generative processes of iteration and recursion. As mentioned earlier, discretization and combinatorial processes are two sides of the same coin—the symbol. We saw that the formation of a shared cache of discrete sounds qua abstract tokens was afforded by the self-organizing interaction between sound-using agents. But this was not the end of the self-organizing multi-agent interaction. With the birth of discretized units also comes the generative combinatorial processes that increasingly furnish such units with syntactic structures.

With the emergence of discrete sounds as transactional tokens, the interactions between agents can further evolve by way of the basic combinatorial processes that hold between discretized symbolic units. It is now the self-organization of language, the syntactic-semantic interaction between agents, that is the real protagonist of our toy universe. The automata are simply agents or players caught up in a game which is the realization and development of language. Any structural change in this multi-agent interaction—language—will translate into a change in the structuring abilities of the automata. This is of course nothing but the reiteration of Hegel's insight that the supposed transcendental subject is only an agent in so far as it is suspended in the process of the self-organization of geist, whose *dasein* or presence is encapsulated in the self-organization of language.

The most generic forms of generative combinatorial processes afforded by discretized sounds are *iteration* and *recursion*. Through these two generative processes, the concatenation of sounds can take on increasingly complex syntactic structures. Let us briefly and intuitively examine what these processes are, how they can structure the automaton's nonconceptual representings of their world, and finally how they construct syntactic vocabularies which the automata can use in their interactions.

Iteration is a memoryless combinatorial process, where constituents can be repeated and combined without restriction. Take for instance the instruction 'chop the garlic until it is turned into a paste'. To yield the desired output (paste), one does not need to memorize the history of the previous actions (choppings). As long as the action of chopping is repeated enough, one will get the paste at the end of the process. In contrast, recursion is a memory-driven generative process, meaning that the performance of each action is built on the history of previous actions. Succinctly formulated, recursion is the embedding of constituents within constituents of the same kind or category.[243] An intuitive example of recursion would be the instruction 'cut the pie into eight equal slices'. The output of this operation cannot be obtained unless each action (cutting to equal slices) is embedded in the history of the previous step. The first cut gives two slices, the second four slices, and finally in this manner, the fourth cut combined with or embedded in previous actions of the same kind yields eight equal slices.

Iteration and recursion can be distinguished according to the types of structure they generate. Simple recursion generates syntactic structures that have complex dependency relations between constituents of the same kind. Owing to the absence of embedding constraints (i.e., the combination of constituents of the same kind), simple iteration generates structures that lack dependency relations but contain constituents of different kinds. The combination of iterative and recursive operations generates structures with hierarchical dependency relations where there are generative rules pertaining to the constituents, and transitions rules between different hierarchies.

In our toy universe, simple iteration of sound-tokens (α, β, γ) would generate syntactic strings such as α, $\alpha\beta$, $\alpha\beta\gamma$, $\alpha\beta\gamma\alpha$, ...

Since iterative operations are memoryless in the sense that each step proceeds independently of the previous one, in their rudimentary form they can generate unlimited concatenations. When iterative operations are combined with embedding operations (i.e., the operation of embedding one constituent in another constituent), they can also encode dependency-relations

243 See H. van der Hulst (ed.), *Recursion and Human Language* (Berlin: de Gruyter Mouton, 2010).

between constituents, as in the case of: $[\alpha]$, $[\alpha[\beta]]$, $[\alpha[\beta[\gamma]]]$, $[\alpha[\beta[\gamma[\alpha]]]]$, ... In contrast to the previous class of strings (α, $\alpha\beta$, ...), these strings contain information regarding dependency relations between constituents. For example, depending on its formal semantic interpretation, the syntactic string of the form $[\alpha[\beta]]$ could mean the right constituent ß is dependent on the left constituent α.

Finally, applying the operation of recursive embedding to sound-tokens generates syntactic strings or expressions capable not only of encoding dependency-relations between constituents, but also of producing new hierarchies between constituents of the same category. Since each recursive step operates on the product of the previous steps (i.e., it is a memory-driven operation), it can embed constituents of the same category within each other according to the history of their past transformations. Using this procedure, recursive embedding produces new hierarchies that encode the history of transformations between constituents of the same category. For this reason, syntactic expressions produced through recursive embedding contain a wide range of coded information regarding the structure, the type of grouping, and their intracategorical dependencies.

The combined application of iterative and recursive operations to the sound-tokens results in the proliferation of syntactic expressions bearing an extremely diverse range of information regarding the encoding details of dependency relations, transformations, hierarchies, and types. In this fashion, syntactic expressions can be grouped based on hierarchies of the complexity of their structure and the information they encode. This is the basic idea behind Chomsky's revolutionary contribution to artificial intelligence in the context of the theory of formal grammar or syntax: From a limited set of discrete units and through the implementation of generative processes, a nested hierarchy of syntactic complexity can be constructed where each level represents a specific class of syntactic language or grammar with its own set of production or computational rules. These syntactic structures can be classified, from top to bottom, into recursively enumerable, context-sensitive, context-free, and regular languages. Moving from the bottom (languages recognized by automata with finite memory) to the top (recursively enumerable languages recognized by automata with infinite