

Title:

From Sensorimotor Sequence to Grammatical Construction: Evidence
from Simulation and Neurophysiology

Author:

Peter Ford Dominey

Institut des Sciences Cognitives

CNRS UMR 5015

67, Boulevard Pinel

69675 BRON Cedex

FRANCE

Telephone: 00 (33) 4 37 91 12 12

Direct line: 00 (33) 4 37 91 12 66

FAX: 00 (33) 4 37 91 12 10

email: dominey@isc.cnrs.fr

WEB: <http://www.isc.cnrs.fr/>

Abstract:

The current research describes a functional trajectory from sensorimotor sequence learning to the learning of grammatical constructions in language. A brief review of the functional neurophysiology of the cortex and basal ganglia will be provided as background for a neural network model of this system in sensorimotor sequence learning. Sequential behavior is then defined in terms of serial, temporal and abstract structure. The resulting neuro-computational framework is demonstrated to account for observed sequence learning behavior. More interestingly, this framework naturally extends to grammatical constructions as form-to-meaning mappings. Predictions from the neuro-computational model concerning parallels in language and cognitive sequence processing are tested against behavioral and neurophysiological observations in humans, resulting in a refinement of the allocation of model functions to subdivisions of Broca's area. From a functional perspective this analysis will provide insight on the relation between the coding structure in human languages, and constraints derived from the underlying neurophysiological computational mechanisms.

1. Introduction:

The human language capability is so distinct from all other behavior, that it has been tempting to explain it with highly specific dedicated modules and innately specified capabilities. Clearly, the species specific language capability indicates that there is a genetically coded language capability. The current research attempts to determine to what extent the language capability is built upon preexisting neurophysiological systems dedicated to the processing of spatiotemporal sensorimotor behavior.

It is worth noting the recent historical context of the problem. In his review (Chomsky 1959) of Skinner's (1957) *Verbal Behavior*, and subsequent essays, Chomsky defined a research program for linguistics in which language was addressed primarily from the perspective of its formal properties. Learning was largely characterized in terms of grammar induction, and the behaviorist emphasis on meaning was significantly reduced. In this context the 'poverty of the stimulus argument' was developed, essentially stating that from the perspective of learning a grammar, the input to the child is highly underspecified - it does not constrain the space of possible grammars. This argument was bolstered by a famous paper from Gold (1967), which provided formal proof supporting the poverty of the stimulus argument, essentially proving that under rigorous constraints of learning success, language cannot be learned by positive evidence alone, and thus requires some alternative method to restrict the learning problem. To address this type of problem, a Universal Grammar was proposed (Chomsky 1965), and language acquisition was defined in terms of determining the UG parameters for the target language based on limited input. Of course, these arguments rely on the somewhat arbitrary characterization of the input in language acquisition, and the characterization of what is being learned, and when. With respect to what is being learned, proponents of UG tend to argue for the 'continuity hypothesis' which

holds that there is a functional continuity between the grammar of the child and that of the adult. The child has access to the adult UG and once a given parameter is set, it becomes available in a generalized manner.

More recently both of these assumptions have been challenged. Research in the developmental psychology of language acquisition has provided strong arguments for the 'richness of the stimulus.' In this context, Tomasello (2003) reviews extensive data indicating the importance of social pragmatic interaction in the acquisition of language. In this framework, far from being a question of formal grammar identification, language acquisition involves the construction of a shared space for attention and meaning that allows the child to learn highly specific language-meaning relations that later become generalized. In this context, grammatical constructions define the functional mappings between the surface form of utterances and the corresponding structural form of the semantic representations of the meaning of these utterances. With respect to the continuity hypothesis, this usage based framework stipulates that the initial form to meaning mappings will be quite specific and non-generalizing, corresponding to the holophrase period attributed both to development and evolution (Wray 2000, Dominey 2004). This corresponds then to a form of 'discontinuity hypothesis' in which knowledge of grammar is first rudimentary, mapping entire utterances to fixed meanings. Subsequently, free arguments are introduced into the constructions, allowing variation in specification of agent, object and recipient roles. Still later, the grammatical forms become truly compositional, an issue that will be addressed in the discussion.

From this perspective, the objective of the current review is to describe how well documented neurophysiological capabilities for the perception and generation of sequential behavior could provide the basis for language processing in this context. A functional hierarchy of sequential behavioral structure will first be presented, along with a description of the

corresponding neurophysiology and simulation background. In the transition from this behavioral sequence taxonomy and language processing, a set of primitive language-related sequencing functions will be identified, along with their simulation. This provides the background for the specification of how these cognitive sequencing capabilities can provide the basis for the expression of adult language comprehension capabilities.

A strong prediction from this functional characterization of language processing in the context of cognitive sequencing is that there is a certain form of functional equivalence between sequence and language processing, as both can be accounted for by the same model. This hypothesis is then validated by results from neuropsychology, event-related brain potential and functional neuroimaging studies that support this 'equivalence hypothesis.' Given the functional success of this characterization, and increasing resolution on the allocation of computational functions to brain regions, the model is updated to reflect this new level of detail.

2. Cognitive Sequencing Hierarchy and a Neural Basis

One of the most prevalent aspects of human behavior is its inextricable embedding in the relentless flow of time. All behavior takes place in time, and thus the management of the temporal structure of perceptual and motor events is of central importance. Likewise, it must be the case that behavioral systems can adapt and learn novel sequential structures, rather than relying on purely fixed patterns. In order to address the ability to learn and manipulate behavioral sequences, it is first of interest to define distinct components of sequential structure and their corresponding neurophysiological correlates. In this context, a fundamental aspect of sequential structure is serial order. From this perspective sequences ABC and ACB are distinct. A second level of sequential organization is temporal or rhythmic structure. From this perspective A---B-C and A-B---C are distinct. The final level of organization that we will

consider has to do with the potential relation between two sequences such as ABCBAC and DEFEDF. While these sequences vary completely in their serial order they share something in common at a more abstract level. They both adhere to a “rule” of the form 123-213 in their construction, and we say that they share a common abstract structure. Note that abstract structure has the interesting property of being generative, in that once the abstract rule is learned it can be used to generate an open set of new sequences (Dominey et al. 1998).

2.1 Encoding Serial Structure in the Cortico-Striatal System

Looking for the neurophysiological correlate of this ability to manage sequential behavior, we encounter two of the most prevalent and fundamental aspects of the organization of the primate brain architecture. The representational structure of neocortex is organized in a posterior to anterior spatial gradient from primary sensory and motor to abstract integrative representation that culminates in the prefrontal cortex (Fuster 1987). One of the principal structural characteristics of the frontal and prefrontal cortices is the abundance of 'recurrent' connections i.e. connections that leave and then reenter the originating cortical area (Goldman-Rakic 1987). From a computational perspective, information that is present at a given instant in time will traverse these recurrent connections and arrive again, later in time. By this mechanism the recurrent connections allow past events to influence the processing of newly arriving sensory events. This ability to take the temporal history of previous events into account is precisely what is required for representing sequential structure. Barone and Joseph (1989) recorded neurons from the dorsolateral prefrontal cortex of monkeys that had been trained to perform a sequence reproduction task, in which they first observed a sequential illumination of lighted push buttons, and were then required to touch the buttons in the remembered sequence. Visual-tonic neurons

encoded both the spatial location of a given target, along with its rank within the sequence. Thus, a given neuron might represent the left target but only when it was first (and not second or last) in the sequence. Dominey et al (1995) developed a neural network simulation of the prefrontal cortex as a network of recurrently connected neurons that precisely reproduced this sequence coding property presented in Figure 1.

Figure 1A represents the detailed neurophysiological model from Dominey et al (1995), and 1B represents the essential functional components of the system referred to as the Temporal Recurrent Network (TRN) from Dominey and Ramus (2000). In both cases, the fundamental capability to encode sequential structure is realized by a recurrently connected network of leaky integrator neurons corresponding to PFC in A and State-StateD in B.

Equations (1.1) and (1.2) describe how the 5x5 unit layer State in Fig1B is influenced by external inputs from Input, recurrent inputs from StateD, and responses from Out. This recurrent state network was modeled after primate frontal cortex with its recurrent corticocortical connections (Goldman-Rakic 1987), and allowed us to explain the electrophysiological encoding of visual space and sequential context (Dominey et al. 1995) recorded in neurons of the primate prefrontal cortex while monkeys performed learned movement sequences (Barone and Joseph 1989). Equation (1.1) describes the leaky integrator, $s(t)$, corresponding to the membrane potential or internal activation of State. In Equation (1.2) the output activity level of State is generated as a sigmoid function, $f(s(t))$, of $s(t)$. The term t is the time, Δt is the simulation time step (5ms), τ is the leaky integrator time constant. As τ increases with respect to Δt , the charge and discharge times for the leaky integrator increase. The unit of time in the simulations is referred to as a simulation time step or sts, and corresponds to a single update cycle of the simulation.

$$s_i(t+\Delta t) = (1 - \frac{\Delta t}{\tau}) s_i(t) + \frac{\Delta t}{\tau} (\sum_{j=1}^n w_{ij}^{IS} \text{Input}_j(t) + \sum_{j=1}^n w_{ij}^{SS} \text{StateD}_j(t) + \sum_{j=1}^n w_{ij}^{OS} \text{Out}_j(t)) \quad (1.1)$$

$$\text{State}(t) = f(s(t)) \quad (1.2)$$

The connections w^{IS} , w^{SS} and w^{OS} define the projections from units in Input, StateD, and Out to State. These connections are one-to-all, and are mixed excitatory and inhibitory, and do not change with learning. This mix of excitatory and inhibitory connections ensures that the State network does not become saturated by excitatory inputs, and also provides a source of diversity in coding the conjunctions and disjunctions of input, output and previous state information.

Recurrent input to State originates from the layer StateD. StateD (Equation 2.1 and 2.2) receives input from State, and its 25 leaky integrator neurons have a distribution of time constants from 20 to 400 simulation time steps, while State units have time constants of 2 simulation time steps. This distribution of time constants in StateD introduces a “damping” (hence the D) or low pass filtering that yields a range of temporal sensitivity similar to that provided by using a distribution of temporal delays (Kühn & van Hemmen 1992).

$$sd_i(t+\Delta t) = (1 - \frac{\Delta t}{\tau}) sd_i(t) + \frac{\Delta t}{\tau} (\text{State}_i(t)) \quad (2.1)$$

$$\text{StateD} = f(sd(t)) \quad (2.2)$$

The model has similarities with previous recurrent models (Elman 1990, Pearlmutter 1995, Pineda 1989) with three important differences. First, there is no learning in the recurrent connections (i.e. those that project from State_D to State and back), only between the State units and the Out units (or PFC and CD in 1A). Second, adaptation is based on a simple associative learning mechanism rather than back-propagation of error, or related error-gradient calculation methods (Pearlmutter 1995). Third, in the temporal domain, a) the computing elements are leaky integrators, and b) simulation time steps are not tightly coupled to input, output and learning processing. That is, an input event can be specified to have a duration of any arbitrary number of times steps, and temporal delays between inputs can likewise be specified. Eliminating the need for calculation of weight changes in the recurrent connections makes this technically and neurophysiologically straightforward. Indeed, the experimenter's capability to specify the time delays between external events is an integral part of this model (Dominey et al. 1995).

The question then is how do these cortical representations of sequence structure become linked to behavior. Here we encounter one of the second fundamental organizational principals of primate (and mammalian) cerebral architecture. With the exception of the most purely primary sensory cortices, essentially the entirety of the neocortex projects in a highly structured and systematic manner onto the striatum - the primary input node of the subcortical basal ganglia (Alexander et al. 1986). The basal ganglia then provide access to the sensorimotor system via their influence on the thalamus. This is illustrated in Figure 1A as prefrontal cortex (PFC) projects onto the caudate nucleus (CD) of the striatum, and in 1B as State projects to Output. The CD then projects to thalamus and the oculomotor output system (the superior colliculus – SC in Figure 1). The question remains, how can this be used for learning behavioral sequences? Such

a mechanism would require the arbitrary binding of sequence state encoding activity in PFC with motor response activity in the caudate nucleus (CD).

In the primate brain, this binding is provided by multiple mechanisms for synaptic plasticity initiated by the influence of the neurotransmitter dopamine in the striatum (Centonze et al. 2001). Extensive investigation of subcortical dopamine producing neurons indicates that their activation (and resulting release of dopamine in the striatum) is triggered by behaviorally induced reward, and stimuli that predict the subsequent arrival of reward (Schultz 2004).

During multiple trials in behavioral sequence learning, the random chance choice of the correct response in a given behavioral context yields the generation of a reward from the environment. This reward triggers the release of dopamine in the striatum (CD in Figure 1A) that in turn activates a series of molecular events that strengthen the synaptic connections between cortical PFC neurons encoding the sequential context, and striatal CD neurons encoding the correct response. This corresponds in Figure 1B to the modifiable synapses between State and Out.

The associative memory in Fig 1Bis implemented in a set of modifiable connections (w^{SO}) between State and Out, described in equation (3). Each time a response is generated in Out, it is evaluated and the connections between units encoding the current state in State, and the unit encoding the current response in Out are modified as a function of their rate of activation and learning rate R . R is positive for correct responses and negative for incorrect responses. Weights are normalized to preserve the total synaptic output weight of each State unit, thus avoiding saturation with extensive learning.

$$w^{SO}_{ij}(t+1) = w^{SO}_{ij}(t) + R * State_i * Out_j \quad (3)$$

The network output is thus directly influenced by the Input, and also by State, via learning in the w^{SO} synapses, as described in Equations (4.1) and (4.2). In Equation (4.2) the sigmoid output function $f' is the same as $f()$ in Equations (1.2) and (2.2), and it additionally performs a winner-take-all function so that only one output neuron is active in the generated response.$

$$o_i(t+\Delta t) = (1 - \frac{\Delta t}{\tau}) o_i(t) + \frac{\Delta t}{\tau} (Input_i(t) + \sum_{j=1}^n w^{SO}_{ij} State_j(t)) \quad (4.1)$$

$$Out = f'(o(t)) \quad (4.2)$$

The model reliably simulated both the sequence learning behavior, and the neurophysiological activity of prefrontal cortex neurons recorded in the behaving primates described in Barone and Joseph (1989). Dominey (1995) subsequently demonstrated the ability of this model to simulate the learning of complex sequences, reinforcing the notion that this recurrent network has inherent capabilities for encoding sequential structure. The model is fairly stable in the face of changes to the fixed parameters. Thus, the time constants in the StateD units can vary by up to 100%, and the temporal delays between successive stimuli by up to 200% before producing significant impairments in sequence learning, depending on the sequence length and complexity (see Dominey 1995).

2.2 Encoding Temporal Structure in the Cortico-Striatal System

As described above, then, the corticostriatal system provides a basis for the learning of serial order in behavioral sequences. As the observation of essentially any behavioral sequence will reveal, serial order is not enough. Intrinsic to the serial order of behavioral sequences is the superposition of temporal or rhythmic structure upon the serial order. This is perhaps most obvious in the rhythmic structure of music and prosodic structure of language. In a series of behavioral sequence learning studies using a serial reaction time (SRT) protocol (Nissen & Bullemer 1987), we exposed human subjects to behavioral sequence learning tasks with sequences such as ABCABDCD in which each element A-D corresponded to a button press on a touch sensitive screen, and sequence learning was measured by the successive reduction in response time to these elements when they occurred in the repeating sequence (Dominey 1998a,b). We modified this standard serial reaction time task by associating specific delays with each stimulus while the sequence was being learned. Simulation studies with the model predicted that under these conditions, if the temporal delay structure was modified between the training and the testing sequence, this would modify the representation of the sequence in the recurrent network, effectively yielding a different sequence than that which was used during training, and thus resulting in impaired performance with the temporally modified sequence (Dominey 1998b). Indeed, these and related predictions concerning the ability to learn and discriminate between different temporal structures were borne out in the subsequent behavioral studies (Dominey 1998a). This indicates that the recurrent prefrontal cortical networks and the plastic cortico-cortical projection provides the basis for the representation and learning of serial and temporal structure.

It is worth noting that the simulated recurrent prefrontal network relies on fixed randomized recurrent connections, with no learning-related modification of these connections. Related recurrent networks (e.g. the Simple Recurrent Network – SRN - Elman 1990) often employ learning on these recurrent synapses, and in order to reduce the computational complexity of this learning, temporal delays are typically compressed into a single network update cycle. For this reason we have referred to the PFC recurrent network as the Temporal Recurrent Network or TRN, based on its ability to process temporal delays in an efficient and realistic manner (Dominey and Ramus 2000, Blanc et al 2003).

2.3 Encoding Abstract Structure in the Cortico-Striatum System

During a series of behavioral experiments with complex sequences of the form ABCDABCEABCFABCGABCH, it was observed that when subjects were exposed to multiple sequences of this form in which the symbols A-D were systematically substituted by different locations on the touch sensitive screen, they began to learn the underlying regular structure and to transfer this learning to new sequences (these unpublished observations were validated in Dominey et al. 1998). This was evidence that sensorimotor sequence learning can benefit from abstract structural correspondences between distinct sensorimotor sequences.

We thus set out to study this abstract learning capability in a systematic manner. This can be characterized as the ability to learn an abstract rule (e.g. ABCBAC) that can then be used to generate new sequences (e.g. HBSBHS). The ability to learn such abstract sequences in which variables could be replaced by classes of percepts or actions would clearly be of value in the sensorimotor domain, as it would allow the generation of novel movement sequences based on learned, well-formed templates. Interestingly, the recurrent network failed in this type of

learning, because it lacked the ability to represent the repetitive structure (see Marcus et al. 1999). Dominey et al (1998) thus modified the corticostriatal model, adding a working memory of the last N elements against which new elements could be compared, illustrated in Figure 2. The neural implementation of a spatial working memory is straightforward, as is that for a comparator or pattern matcher that determines whether two input vectors are identical or not. With this working memory and a simple pattern matcher, the sequence ABCBAC was represented as 'u, u, u, -2, -4, -3' where 'u' indicates a *unique* or *unmatching* element, -N indicates an element that matches with the element N positions back in memory. This abstract coding is then represented in the recurrent network, which thus encodes the abstract structure of the presented sequences. The idea then is that while serial and temporal structure are encoded in the corticostriatal system, abstract structure requires additional, dissociated coding capabilities.

In particular, the working memory function will require a neural system that allows the representation of arbitrary elements within a particular ordered configuration in order to subsequently manipulate that configuration. This is functionally similar to the ability to mnemonically represent arbitrary visual objects at specific spatial locations for subsequent spatial manipulation. Indeed, based on the distinction between visual processing systems for object properties (the ventral system) vs. spatial properties (the dorsal system) Ungerleider et al. (1998) have extended this neurophysiological distinction to prefrontal cortical memory areas, with the dorsal spatial processing stream projecting to prefrontal areas adjacent to and partially overlapping with Broca's area. In the Abstract Recurrent Network, we can consider that the working memory function is thus performed by these prefrontal spatial working memory areas (see Chang 2002).

3. Transition to Language

Given these capabilities for serial, temporal and abstract structure processing, we can now consider how these capabilities can contribute to language processing. In this context, Dominey and Ramus (2000) first demonstrated that the dual process ATRN model could account for young infants abilities to process the serial, temporal and abstract structure of language-related sound sequences as revealed in documented experimental studies. Clearly, however, it would be more interesting to demonstrate how the model could account for more adult-like language behavior.

This lead to the non-trivial problem choosing a behavioral measure of language processing that could be tested on the model. Within the context of the “formalist” vs. “functionalist” continuum, we were more interested in a test that would address aspects of language including comprehension and meaning, rather than purely formal aspects. A survey of the neuropsychological literature revealed a behavioral language task that was well suited for adaptation to the model in this context. Caplan et al (1985) developed a behavioral protocol to assess brain-lesioned patients’ ability to determine “who did what to whom” based purely on the syntactic structure of the target sentences. They presented subjects with sentences constructed from a set of nine different grammatical construction types, and for each sentence, asked the subject to respond by indicating the agent, object, and recipient (always in that canonical order) of the action described in the sentence. This thematic role assignment was performed by the subjects via their pointing to photographs in the required agent, object, and recipient order. Thus, the input was a spoken sentence and the output was the behavioral sequence of pointing to agent,

object and recipient, always in this 'canonical' order. In this context, consider the following two sentences:

(1) John gave the ball to Mary.

“AGENT *verb* OBJECT to RECIPIENT”; < AGENT, OBJECT, RECIPIENT >

(2) The ball was given to Mary by John.

“OBJECT was *verb* to RECIPIENT by AGENT”; < AGENT, OBJECT, RECIPIENT >

For both of these sentences the agent, object, recipient are respectively John, Ball, and Mary. In the behavioral task in which the subject should indicate the agent, object and recipient, in that order for sentence (1), this response sequence can be generated simply by repeating the nouns in the same order in which they occurred in the input. We can thus characterize the structure of the transformation from input to output as AOR-AOR, or 123-123. Alternatively, in the analogous type of analysis for sentence (2) the input noun order does not correspond to the required agent, object, recipient response, and thus must undergo a structural transformation corresponding to ORA-AOR or 123-312. Viewed in this manner, the thematic role assignment task of Caplan et al (1985) can be reformulated behaviorally as an abstract cognitive sequencing task, in which subjects must learn the appropriate abstract structure for each of the nine different grammatical constructions used in the task.

From the perspective of the model, this abstract structure manipulation is part of its built in capability. The only remaining problem is to allow the system to learn to recognize and correctly manipulate multiple abstract structures, each corresponding to a distinct grammatical construction. In this context, Bates et al (1982) suggested that across languages, grammatical constructions are identified by a highly restricted set of cues including word order, word

category, grammatical function words, and prosodic structure. Restricting the analysis to grammatical function words, it can be seen that in (1) and (2) above, if the open class words (i.e. the nouns and verbs) are considered as fillers for variables 'slots' then the configuration of remaining grammatical function words forms a sequence that uniquely identifies the corresponding grammatical constructions. The grammatical construction frames corresponding to sentences (1) and (2) are thus illustrated. In other words, different construction types can be identified by their configuration of grammatical function words (and or bound morphemes). This can then be used as an index to store and retrieve the corresponding form to meaning mapping.

Based on this analysis, Dominey (2001) and Dominey et al (2003) used the dual process model such that grammatical function words were processed in the recurrent network, and open class words were processed in the short term working memory. Through training, the model learned to associate the pattern of function words encoded in the recurrent network with the appropriate corresponding abstract transformation structure. In this manner the model was able to learn the nine grammatical constructions that were tested in the Caplan test of syntactic comprehension.

This raises the issue of how the lexical categorization (i.e. the separation of open and closed class words) could be achieved. It has long been observed that across languages, grammatical function words (such as determiners and prepositions) are generally shorter and less stressed than content words (e.g. Shi et al. 1999). Recently, Shi et al (1999) demonstrated that human infants can reliably perform this perceptual categorization of open vs closed class words based on these words acoustic properties. In this context, Blanc et al (2003) formalized such observations with the proposition that the presence of peaks in the pitch or fundamental frequency (F0) of the acoustic signal (reflecting in part the presence of accent or stress) would predict that the associated word is an open class word. Classification based on this F0 peak

detection proved reliable, but left open the question of the underlying neuro-computational mechanisms. In order to respond to this question, Blanc et al (2003) tested the TRN in this form of temporal discrimination task based on F0 structure and observed, indeed, that the TRN reliably performed the lexical categorization task.

4. Neurophysiology and Model Validation

Figure 3 represents the transition from the TRN network to the ATRN network, and the functional equivalence between abstract sequence processing and thematic role assignment in syntactic comprehension. The corresponding analysis outlined above makes a rather remarkable statement about the underlying neurophysiology of at least certain stereotypical aspects of sentence processing. In particular, it implicitly states a form of 'equivalence hypothesis' which holds that the same neural systems invoked by the processing of nonlinguistic behavioral sequences that possess the appropriate degree of abstract structure are also responsible for the thematic role processing as functionally characterized above. The following sections present data from neuropsychological and brain imagery tasks that test the predictions of this "equivalence hypothesis", and provide additional data for further elaboration of the model.

4.1 Aphasics

The first effort to test this equivalence hypothesis was undertaken with populations of neurological patients that demonstrated syntactic comprehension deficits as revealed by their poor performance on the Caplan (et al. 1985) task described above. The equivalence hypothesis

predicts that these patients' impairments on the Caplan task would be correlated with their impairments on a test of abstract sequence structure processing. Lelekov et al (2000b) thus tested patients with lesions in and around Broca's area on the Caplan task, and a task in which they were required to learn the abstract structure ABCBAC and then judge whether 20 new sequences corresponded to this abstract structure or not. As illustrated in Figure 3B and 3C, we predicted that these two tasks could be realized by a shared neurophysiological mechanism. As predicted, the impairments in the two tasks were highly correlated, arguing in favor of a common underlying mechanism (Lelekov et al. 2000b, Dominey et al. 2003).

This type of correlation does not necessarily demonstrate a common mechanism, however, and a stronger argument could be made if it were demonstrated that training in one of the domains led to a transfer of performance to the other. We thus set out to achieve this demonstration by training aphasic patients on an abstract non-linguistic sequence whose transformation structure corresponds to that required for transforming a complex sentence type into the simple canonical order AGENT VERB OBJECT. Transfer of improved performance from the sequence domain to the sentence processing domain would provide further evidence for a common underlying mechanism. Thus, Hoen et al (2002) trained agrammatic aphasic patients on sequences of the abstract structure ABC-BCA that corresponds to the transformation of relativised sentences 'It was the ball that John took' to 'John took the ball' (Ball John Took – John Took Ball). The patients were evaluated on the Caplan task before and after 10 weekly sessions of training on nonlinguistic sequences generated from the ABCBCA structure. The patients displayed a reliable capability to learn the abstract structure, and most interestingly a reliable transfer of this training only to the corresponding relativised sentences, and not to other grammatical construction types. The observed transfer from sequence learning to sentence

comprehension, specific to the target grammatical structure, argues in favor of the equivalence hypothesis.

4.2 Event Related Scalp Potentials

These behavioral correlation data suggest a common underlying mechanism. In order to investigate this further, Hoen and Dominey (2000) compared event related brain potentials (ERPs) evoked during the processing of sentences and abstract sequences. In ERP studies, an array of electrodes on the scalp measures electrical current on the scalp generated by active populations of cortical neurons that form electrical dipoles. Multiple successive trials are averaged in order to provide a clear electrical signal, thus these studies require the identification of specific behavioral events on which to align the recordings for averaging. In sentence processing, recordings are thus typically aligned on word onset for successive words. A number of ERP studies have demonstrated significant differences in brain activity in response to grammatical function words vs. content words. Content words tend to evoke a central negative effect at about 400ms after word onset, the N400. In contrast, grammatical function words tend to evoke a left anterior negativity (LAN) during a period 400-600 after word onset (Brown et al. 1999).

Grammatical function words indicate the subsequent grammatical sentence structure and potential transformation processing required for thematic role assignment. A parallel can be drawn this role, and the role of “function symbols” such as *X* and *Y* that indicate different transformation processing in sequences ABCXABC and ABCYBAC, respectively. Of particular interest was the comparison between grammatical function words in sentences, and function symbols such as *Y* in the sequence ABCYBAC during a correctness judgement task. In this task,

subjects learned that *Y* indeed played the role of an abstract function word, signaling the upcoming transformation of the initial triplet. Both in natural language sentences and in abstract sequences, these function symbols evoked a left anterior negativity (LAN) in the 400-600 ms timeframe that was essentially indistinguishable in these two conditions. Likewise, in both cases this LAN effect for function words or symbols was significantly different from the N400 effect for content words and symbols, respectively (Hoen & Dominey 2004).

These studies indeed indicate a partial overlap in the functional neuroanatomy of sentence comprehension and abstract sequence processing. Clearly, however, there are obvious contrasts between these behaviors, and the corresponding neurophysiological differences are of interest. Such differences should be revealed using a finer grain of spatial localization.

4.3 Functional MRI

In this context, Hoen et al (2005) have further investigated the underlying neurophysiology in a functional MRI study comparing language and sequence processing. Part of the objectives of this work was to identify not only the common neurophysiological processes, but equally importantly to identify those processes that differentiate between language and sequence processing. Based on the simulation studies, we would predict that sentence and abstract sequence processing would activate a common brain network functionally associated with the use of the transformation working memory, and the recurrent network for processing the function elements depicted in Figure 2. In addition to this network, language processing would lead to activation of brain regions associated with the insertion of word meanings into this transformation processing as required for thematic role assignment.

Interestingly, within the cortical region of Broca's area this theoretical prediction was partially resolved, as illustrated in the model in Figure 4. Subjects were required to read abstract sequences and sentences (in separate blocks of trials) and to judge whether the stimuli were structurally/grammatically well formed. Both during sequence and sentence processing, the pars opercularis (Brodmann's area 44) and adjacent areas 46 and 6 were significantly activated, while the neighboring BA 45 was activated only for sentence processing (Hoen et al. 2005). Neuroanatomical studies have confirmed that the cortical network including areas 44, 46 and 6 overlaps with the frontal projection sites of the dorsal visual stream for spatial working memory (Ungerleider et al. 1998). As mentioned above, the dorsal stream processing of spatial relations between arbitrary objects is functionally analogous to the processing of abstract transformation structures instantiated with arbitrary sequence elements (see Chang 2000). In contrast, area 45 corresponds to the frontal projection site of the ventral object processing visual stream, and implements a working memory capability for object properties. This is coherent within the context of the model, as this corresponds to the integration of semantic object properties into the structure mapping process for the assignment of thematic roles. Thus the model of Figure 4 represents the “evolution” of the model in Figure 2. One principle change is the introduction of this lexical semantics or word meaning into the transformation processing mechanism, carried out by BA 45. A second change is the representation of the processing of closed class words in Construction Index which plays the same role as the recurrent network in the model in Figure 2. The model retains the essential property that the constellation of closed class words or function words is used as an index (the Construction Index) into a structured inventory (the Construction Inventory) of mappings (Transformation) from the input order of open class elements onto the canonically ordered arguments of the meaning representation.

5. Co-Influence Between Neurophysiological Mechanisms and the Structure of Language

What is the potential effect on language of this cerebral organization? One of the principal differences between the prefrontal cortex of non-human primates and humans is the displacement of cortical areas for spatial and non-spatial working memory in man respectively above and below their locations in the monkey brain, making room for new areas in the posterior perisylvian cortex, including Broca's area, associated with language processing (Ungerleider et al. 1998). Within this setting of cortical evolution, it is thus of great interest that functionally, our model of sentence processing relies on specialized forms of working memory in order to perform the transformation from a sentence to a representation of meaning. One of these working memory functions, linked with activation of Brodmann's area 45 and seen only in sentence (not sequence) processing, is associated with the semantic value of words and the integration of this meaning into a mapping mechanism. This is consistent with the characterization of BA 45 as the frontal working memory component of the semantic "object" stream. In contrast, the working memory associated with the actual transformation process appears to activate Brodmann's areas 44 and 46, for transformation processing in sentences as well as non-linguistic sequences. Again, consistent with characterization of this region as involved in spatial "structural" working memory. Thus, the resulting system allows the use of systematic transformations of grammatical structures onto meaning structures, guided by special percept cues – corresponding to function words (and the general class of such cues evoked by Bates and MacWhinney) in language.

As noted above, there is a vast richness across languages with respect to the variety of different mappings of grammatical structures (e.g. active, passive, etc.) onto meaning representations. All languages must provide a method for encoding these possible mappings for each sentence type. Likewise, the brain must be capable of adapting to all of the possible strategies employed within natural languages. The intersection of these two sets of constraints

(1) on how meaning can be encoded, and (2) on the functional organization of the code reading machine, results in the set of constraints that define natural languages. Bates and MacWhinney proposed that the punctuation of the code must be based either on word order regularities (e.g. the first noun is always the agent, the second the object etc.) or on the use of special markers (e.g. grammatical function words like “the” and “by”, or attached grammatical morphemes, or prosodic markers) that indicate the functional roles of associated words or groups of words, and the possible combination of these strategies. Simulation and neurophysiological results indicate that there are specialized neurophysiological structures for the realization of these functions. It is thus interesting to note the close correspondence between these characteristics of natural languages that were derived from extensive cross-linguistic studies (e.g. Bates et al. 1982), and the computing capabilities of the corresponding neural machinery described here.

6. Discussion and Conclusions

The stated objective of this study was to present a trajectory leading from a characterization of the neurophysiology of behavioral sequence learning to a corresponding characterization of the learning and use of grammatical constructions. In this context, a neurophysiologically driven model of cortex and basal ganglia in sequence learning was described, along with its sequence processing capabilities. The model was then extended to address abstract rules for sequence generation. This functional extension provided the essential mechanisms of structure mapping required for making the transition to grammatical constructions as form-to-meaning mappings. Behavioral and brain imagery evidence supporting the proposed model was reviewed and used to refine the allocation of brain areas to computational functions in the model. With respect to insights concerning relations between evolution and acquisition of language, the simulation studies reviewed in this paper provide a characterization of the functional requirements that must

be accounted for in both cases. Current and future research will address how this model captures interesting aspects of the notion of grammatical constructions (Dominey & Boucher 2004), and how this can be of value in simulation and robotic studies of language acquisition (Dominey & Boucher 2005).

As presented, the story leaves at least two important questions open. First – it is important in this model to have a predicate-argument form of representation of meaning onto which sentence structure can be mapped. Recent work on the perception of action by human and non-human primates has revealed the existence of neural systems that represent action – independent of the agent (the “mirror neuron” system) (Rizzolatti & Arbib 1998). Combined with proprioceptive information that allows the differentiation between self and other’s movement, the neurophysiological basis for event representation thus appears well founded. In addition, Hurford (2003) has provided convincing evidence for the neural basis of predicate-argument representations.

The second issue has to do with compositionality. The advantage of learning a grammar is that this knowledge can be used to generate and understand new grammatical constructions that have not previously been experienced. The current implementation of the construction grammar approach has the limitation that a sentence generated from a given grammatical construction cannot be understood until the system has been exposed to a well formed <sentence, meaning> pair using that construction, in order to acquire the appropriate mapping. Interestingly, this limitation appears to mimic the same limitation observed in human language development. It appears that young children during their second year largely limit their constructions to those that have been in speech they have heard (Clark 2003, Tomasello 2003), and the use of such a strategy should take the initial language learner quite some way on the path to adult language capability. Thereafter, the application of pattern finding mechanisms can begin to detect the

grammatical marking of phrase structure (e.g. “that” and relative clauses), so that new constructions can be generated by replacing simple nouns with noun phrases. Thus “The book is on the table” can lead to “The book that we read is on the table”.

In conclusion, the current research attempts to present a coherent trajectory starting with the functional neurophysiology of sensorimotor sequence learning, and leading to the use of grammatical constructions in language. Inherent in this approach is the idea that the principal source of the basic initial structure in language derives from the structure inherent in the conceptual representations that preceded language (Dominey 2002; Jackendoff 1999, 2002). Language then corresponds to the mapping between these conceptual structures and the linear sequences of words in sentences. This mapping capability in turn derives from structural mapping capabilities that are of clear value in the sensorimotor domain. The ability to use abstract sequences like ABCBAC and then fill in A-C with specific sensorimotor components (e.g. objects that are to be manipulated) provides a powerful mechanism for generating novel movement sequences that are invariant to changes in the objects being manipulated while preserving the relations between them. In this context, language remains a uniquely human capability. The point to be made is that a significant portion of this human capability relies on neurophysiologically grounded information processing capabilities that are likely specific neither to humans nor to language.

Acknowledgements: This work was supported in part by funding from HFSP MCILA, ACI Integrative and Computational Neuroscience Initiative, ACI TTT, and the ESF Eurocores OMLL Project.

References:

- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Rev Neurosci* 9 :357-381
- Barone P, Joseph JP (1989) Prefrontal cortex and spatial sequencing in macaque monkey. *Experimental Brain Research*, 78, 447-464
- Bates E, McNew S, MacWhinney B, Devescovi A, Smith S (1982) Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11, 245-299.
- Blanc, J.M., Dodane, C., Dominey, P.F.(2003). Temporal Processing for Syntax Acquisition: A simulation study. 25th Annual Meeting of the Cognitive Science Society. Boston (Massachusetts), USA.
- Brown CM, Hagoort P, ter Keurs M (1999) *J Cogn Neurosci*. May;11(3):261-81.
Electrophysiological signatures of visual lexical processing: open- and closed-class words.
- Caplan D, Baker C, Dehaut F (1985) Syntactic determinants of sentence comprehension in aphasia. *Cognition*, 21, 117-175,
- Centonze D, Picconi B, Gubellini P, Bernardi G, Calabresi P. (2001) Dopaminergic control of synaptic plasticity in the dorsal striatum. *Eur J Neurosci*. Mar;13(6):1071-7
- Chang F (2002) Symbolically speaking: a connectionist model of sentence production. *Cognitive Science* (93) 1 – 43.
- Chomsky N (1959) A Review of B. F. Skinner's *Verbal Behavior*" in *Language*, 35, No. 1, 26-58.
- Chomsky N. (1965) *Aspects of the theory of syntax*. MIT Press
- Clark EV (2003) *First language acquisition*. Cambridge University Press
- Dominey PF (1995) Complex Sensory-Motor Sequence Learning Based on Recurrent State-Representation and Reinforcement Learning, *Biological Cybernetics*, 73, 265-274.

- Dominey PF (1998a) A shared system for learning serial and temporal structure of sensori-motor sequences? Evidence from simulation and human experiments. *Cognitive Brain Research*, 6, 163-174
- Dominey PF (1998b) Influences of Temporal Organization on Transfer in Sequence Learning: Comments on Stadler (1995) and Curran and Keele (1993) *J. Exp Psychology: Learning, Memory and Cognition*, 24 (1) 234-248
- Dominey PF (2000) Evolution of dissociable processing of function vs content words for syntactic comprehension. *Evol Lang 2000*, Paris.
- Dominey PF (2000b) Conceptual Grounding in Simulation Studies of Language Acquisition, *Evolution of Communication*, 4(1), 57-85.
- Dominey PF (2001) A model of learning syntactic comprehension for natural and artificial grammars: in (Witruk, Friederici & Lachmann, eds) *Basic mechanisms of language and language disorders*, Kluwer Academic Publishers, In press.
- Dominey PF (2002) Requirements on Conceptual Representation for the Evolution of Language, *Evol Lang 2002*, Boston.
- Dominey PF (2003) Learning Grammatical Constructions in a Miniature Language from Narrated Video Events, Proceedings of the 25th Annual Meeting of the Cognitive Science Society, Boston
- Dominey PF (2004) From Holophrases to Abstract Grammatical Constructions in Development and Evolution, *Evol Lang 2004*, Leipzig.
- Dominey PF, Arbib MA, Joseph JP (1995) A Model of Cortico-Striatal Plasticity for Learning Oculomotor Associations and Sequences, *J Cog Neuroscience*, 7:3, 311-336
- Dominey PF, Boucher JD (2005) Developmental Stages of Perception and Language Acquisition in a Perceptually Grounded Robot, *Cognitive Systems Research*, In Press

- Dominey PF, Hoen M, Lelekov T, Blanc JM (2003) Neurological basis of language in sequential cognition: Evidence from simulation, aphasia and erp studies, *Brain and Language*, 86 (2003) 207–225
- Dominey PF, Inui T (2004) A Developmental Model of Syntax Acquisition in the Construction Grammar Framework with Cross-Linguistic Validation in English and Japanese, Proceedings of the CoLing Workshop on Psycho-Computational Models of Language Acquisition, Geneva, pp. 33-40.
- Dominey PF, Lelekov T, Ventre-Dominey J, Jeannerod M (1998) Dissociable Processes for Learning the surface and abstract structure sensorimotor sequences. *Journal of Cognitive Neuroscience*, 10 :6 734-751
- Dominey PF, Ramus F (2000) Neural network processing of natural language: I. Sensitivity to serial, temporal and abstract structure of language in the infant. *Language and Cognitive Processes*, 15(1) 87-127
- Elman J (1990) Finding structure in time, *Cognitive Science*, 14:179-211
- Fuster JM (1997) The prefrontal cortex. 3rd Ed. Philadelphia, Lippincott-Raven.
- Gold EM. (1967) Language identification in the limit. *Informat. Control* **10**, 447-474
- Goldberg, A. (1995) *Constructions*. University Chicago Press, Chicago and London.
- Goldman-Rakic PS (1987) Circuitry of the primate prefrontal cortex and regulation of behavior by representational memory. In V.B. Mountcastle (Ed.) *Handbook of Physiology*, 5 373-417.
- Hoen M, Dominey PF (2000) ERP analysis of cognitive sequencing: a left anterior negativity related to structural transformation processing. *Neuroreport* 11(14):3187-91.

- Hoeh M, Dominey PF (2000) Chapter 16 Evidence for a Shared Mechanism in Linguistic and Nonlinguistic Sequence Processing? ERP Recordings of On-Line Function- and Content-Information Integration
- Hoeh M, Pachot-Clouard M, Segebarth C, Dominey PF (2005) When Broca experiences the Janus syndrome. An er-fMRI study comparing sentence comprehension and cognitive sequence processing. *Cortex*, In Press
- Hurford JR. (2003) The neural basis of predicate-argument structure. *Behav Brain Sci.* Jun;26(3):261-83
- Jackendoff R (1999) Parallel constraint-based generative theories of language. *Trends Cogn Sci.* 1999 Oct;3(10):393-400.
- Jackendoff R (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford University Press.
- Kühn R, van Hemmen JL (1992) Temporal Association, in *Physics of Neural Networks*, (Eds E. Domanay, JL van Hemmen and K Schulten), Springer-Verlag, Berlin, 213-280
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77–80.
- Miikkulainen R (1996) Subsymbolic case-role analysis of sentences with embedded clauses. *Cognitive Science*, 20: 47-73.
- Nissen MJ, Bullemer P (1987) Attentional requirements of learning: Evidence from performance measures, *Cogn. Psychol.* 19 1–32.
- Pearlmutter BA (1995) Gradient calculation for dynamic recurrent neural networks: A survey. *IEEE TRANs Neural Networks*, 6(5) 1212-1228.
- Pineda FJ (1989) Recurrent Backpropagation and the Dynamical Approach to Adaptive Neural Computation. *Neural Computation*, 1, 161-172.

- Rizzolatti G, Arbib MA. (1998) Language within our grasp, *Trends Neurosci.* May;21(5):188-94.
- Schultz W (2004) Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr Opin Neurobiol.* Apr;14(2):139-47
- Shi R., Werker J.F., Morgan J.L. (1999) Newborn infants' sensitivity to perceptual cues to lexical and grammatical words, *Cognition*, Volume 72, Issue 2, B11-B21.
- Skinner BF (1957) *Verbal Behavior*, Reprinted by the B. F. Skinner Foundation in 1992 and 2002
- Steels L, Kaplan F (2001) Aibo's first words : The social learning of language and meaning. *Evolution of Communication*, 4(1) : 3 – 32.
- Stolcke A, Omohundro SM (1994) Inducing probabilistic grammars by Bayesian model merging, In *Grammatical Inference and Applications: Proc. 2nd Intl. Colloq. On Grammatical Inference*, Springer Verlag.
- Tomasello, M. (2003) *Constructing a language: A usage-based theory of language acquisition.* Harvard University Press, Cambridge.
- Ungerleider LG, Courtney SM, Haxby JV. A neural system for human visual working memory. *Proc Natl Acad Sci U S A.* 95(3):883-90. 1998
- Wray A (2000) A protolanguage with no declaratives and no names. *Evol Lang 2000*, Paris.

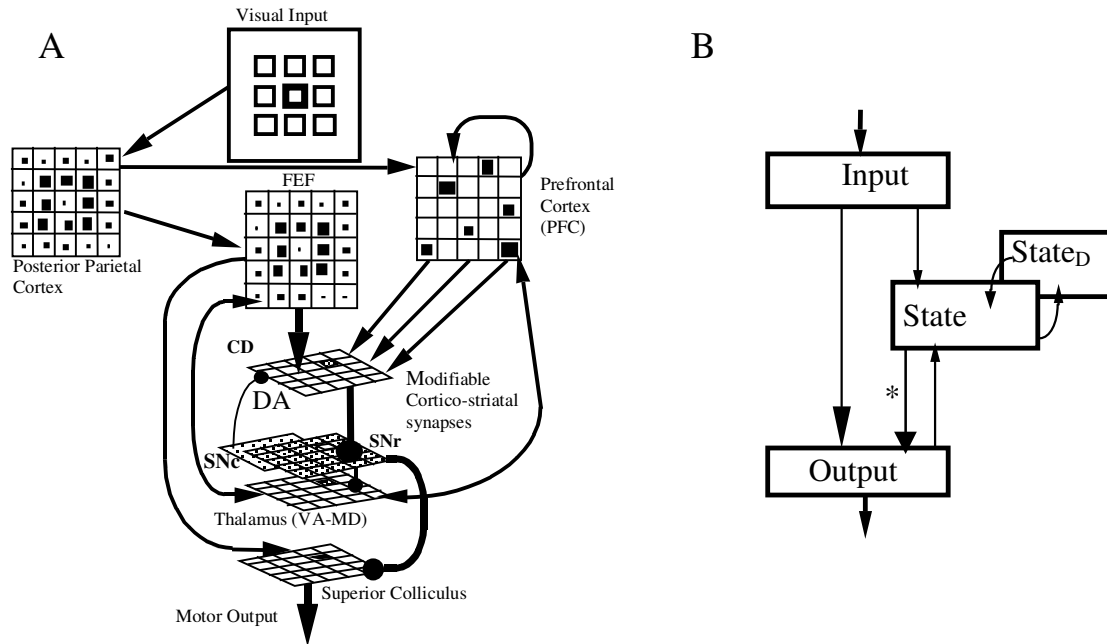


Figure 1. A. The model is based on the functional neuroanatomy of the primate corticostriatal system (Dominey, Arbib & Joseph 1995). Visual inputs traverse the retina and visual system to provide a retinotopic representation of the visual input in Posterior Parietal Cortex and the Frontal Eye Fields (FEF). A non-topographic signal projects from PP to the Prefrontal Cortex (PFC). For movement production, the tonic inhibition of the superior colliculus (SC) from substantia nigra pars reticulata (SNr) is temporarily inhibited by the GABAergic (inhibitory) input to SNr from the caudate nucleus (CD) of the striatum. Caudate cells that code motor outputs are influenced by topographic inputs from FEF and by modifiable non-topographic inputs from PFC. The PFC is a dynamic recurrent network that encodes sequence context. In a sequence learning task, when the correct response is generated, reward related activity of the SNc dopamine (DA) neurons strengthens the PFC-CD synapses that were active in generating the correct response. The caudate thus provides a movement selection function. In this manner, the model can learn context dependant behavior including sequence discrimination and reproduction. B. Simplified version of the model where the recurrent network is explicitly represented as two recurrently connected layers State and State_D (corresponding to PFC), and learning occurs in an associative memory linking State and Output (corresponding to the modifiable connections between PFC and CD in A).

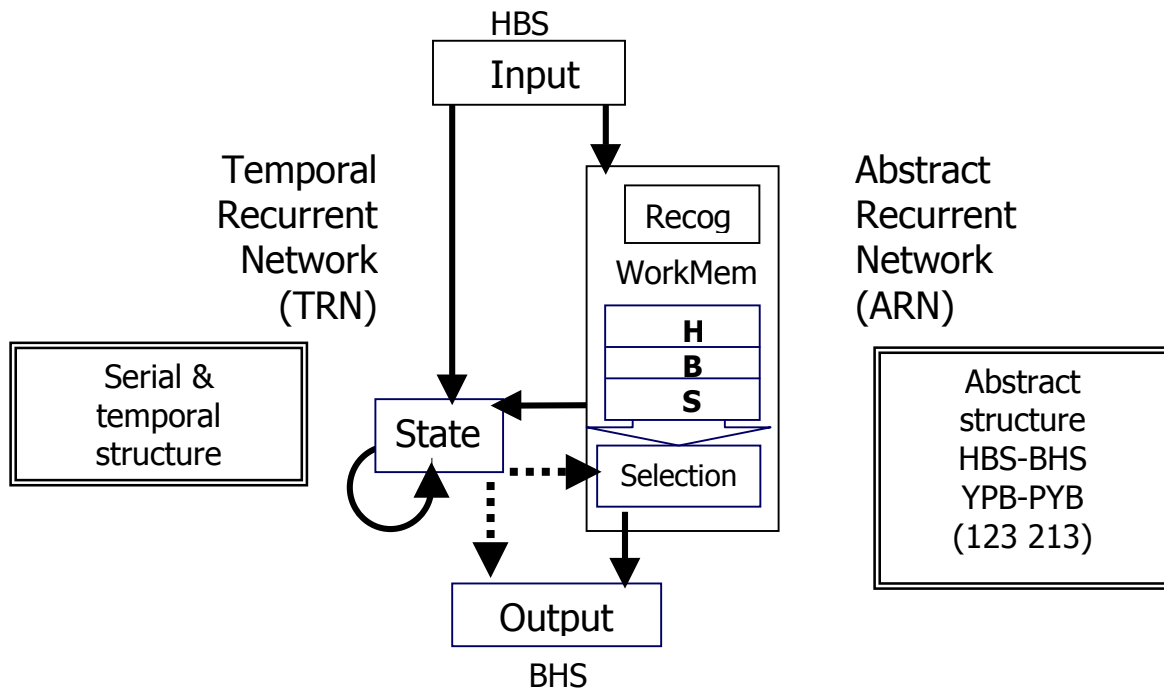


Figure 2. Abstract Temporal Recurrent Network (ATRN) model. The temporal recurrent network (TRN), corresponding to a schematized version of the model in Figure 1, exploits recurrent network (i.e. prefrontal cortex) dynamics to encode serial and temporal structure. To encode the abstract structure common to isomorphic sequences such as HBSBHS and YPBPYB, the abstract recurrent network (ARN) stores the N previous elements of the current sequence in a working memory. The Recognition function compares the current sequence element to working memory contents to detect the abstract repetitive structure. The abstracted coding is represented in the recurrent State network. In the expression of abstract structure knowledge, the contents of the working memory are selectively extracted into the output stream by the activation of selection neurons by State neurons.

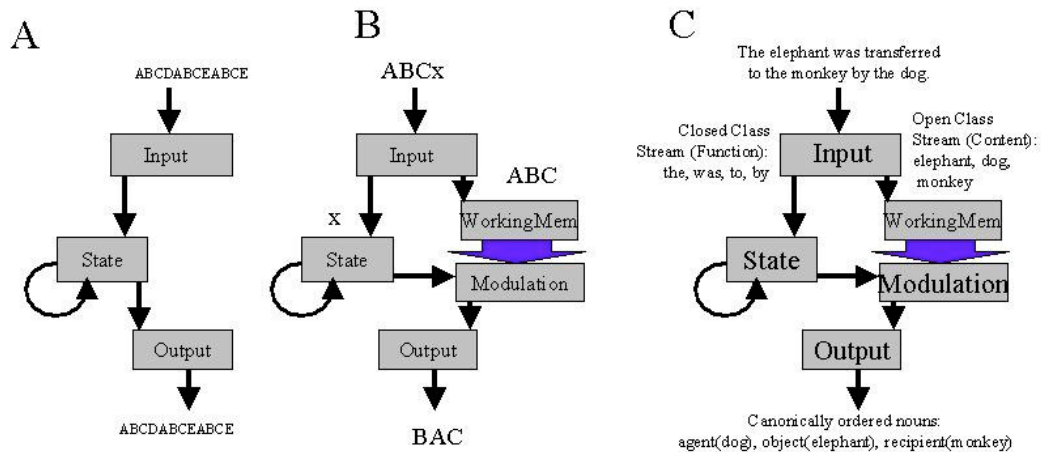


Figure 3. Transition from sequence processing to grammatical constructions. A. Temporal recurrent network. Recurrent network State encodes serial and temporal structure, and influences behavior via modifiable State-Output connections. B. Addition of a Working Memory capability storage of ordered input sequence for recognition of repetitive transformational structure. Abstract structure rules are associated with the sequence of function elements encoded in State. C. Use of the abstract sequence processing model to perform thematic role assignment. See text for explanations.

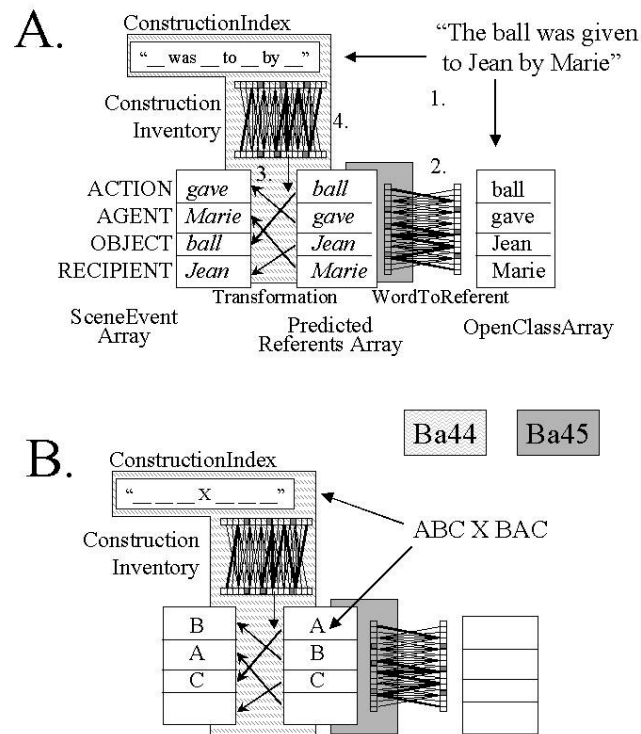


Figure 4. Structure-Mapping Architecture. A. Sentence Comprehension: 1. Lexical categorization - Open and closed class words processed in separate streams. 2. Open class words in Open Class Array are translated to their referent meanings via the WordToReferent mapping. Insertion of this referent semantic content into the Predicted Referents Array (PRA) is realized in pars triangularis Ba45. 3. PRA elements are mapped onto their roles in the SceneEventArray by the Transformation mapping, specific to each sentence type. 4. This mapping is retrieved from Construction Inventory, via the ConstructionIndex that encodes the closed class words that characterize each grammatical construction type. The structure mapping process is associated with activation of pars opercularis Ba44. B. Abstract Sequence Processing: Lexical categorization takes place for function and content elements of non-linguistic sequences (see Hoen & Dominey 2000). As with sentences, function elements allow retrieval of learned transformation from ConstructionInventory via ConstructionIndex.