

Neurological basis of language and sequential cognition: Evidence from simulation, aphasia, and ERP studies

Peter F. Dominey,^{*} Michel Hoen, Jean-Marc Blanc,
and Taïssia Lelekov-Boissard

*Sequential Cognition and Language Group, Institut des Sciences Cognitives, CNRS UMR 5015,
67, Blvd Pinel, 69675 Bron Cedex, France*

Accepted 23 July 2002

Abstract

The current research addresses the hypothesis that certain aspects of sequential cognition have made substantial contributions to the human language processing capability, from a functional neurophysiology perspective. We first describe a cognitive sequence processing model that was developed based on the functional neuroanatomy of primate cortex and basal ganglia. We demonstrate how this model is capable of simulating the behavior of human infants in extracting serial, temporal and abstract structure from language-like sound sequences as revealed in recent psycholinguistic experiments. We then demonstrate how, through training, this model can perform adult level syntactic comprehension, based on dissociated processing streams for open vs. closed class words. The model subsequently predicts: (1) that impaired syntactic processing (as in agrammatic aphasia) will be associated with impairments in corresponding non-linguistic cognitive sequencing tasks, and (2) that neurophysiological processes (as revealed by ERPs) involved in syntactic processing should also be involved in the corresponding non-linguistic cognitive sequencing tasks. Data confirming these predictions are reviewed. We conclude that the study of sequential cognition will provide a new paradigm for the investigation of the neurophysiological bases of language.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Aphasia; Agrammatism; Neural network; Model; Simulation; ERP language

1. Introduction

Understanding the functional neurophysiology of human language remains one of the largest open problems in cognitive neuroscience. The current research addresses this issue by attempting to establish the functional relations between

^{*} Corresponding author. Fax: +043-791-1210.

E-mail address: dominey@isc.cnrs.fr (P.F. Dominey).

sequential cognition and language. Sequential cognition encompasses the capabilities to extract and utilize the sequential structure of perceptual and motor events in the world in an adaptive and pragmatic manner. The fundamental anchoring of all behavior in time places a primordial importance on this cognitive sequencing capability. The generation of goal-directed movements, for example in grasping an object that is partially occluded behind a solid obstacle, requires the generation of a well formed, precise motor command sequence that itself requires transformations between perceptual and motor spaces according to learned, and inviolable transformational rules. In this same sense language also requires the learning and use of transformational rules in the processing and generation of well-formed sentences. The goal of the current research is thus to test the hypothesis that certain aspects of sequential cognition have made substantial contributions to the human language processing capability (Dominey, 1997).

A first step in this direction is the recognition that sequential structure—in language or perceptual motor control—is not unitary, but rather of a multidimensional nature. We consider a division into the serial, temporal and abstract dimensions. Serial structure or order is defined by the relation between an element or set of elements, and its successor. This dimension can be characterized in terms of length and complexity. Temporal structure is defined in terms of the durations of elements (and the possible pauses that separate them), and intuitively corresponds to the familiar notion of rhythm. Thus, two sequences may have identical serial structure and different temporal structure, or the opposite. Abstract structure is defined in terms of generative rules that describe relations between repeating elements within a sequence. Thus, the two sequences ABCBAC and DEFEDF have different serial structure, but are both generated from the same abstract structure 123-213, and are thus said to be isomorphic. An important characteristic of abstract structure is that it can generalize to new isomorphic sequences. While perhaps not exhaustive, these three dimensions at least partially span the space of possible behavioral sequence structure.

These three dimensions of sequence processing have been explicitly demonstrated in human infants in the domain of language processing. Saffran, Aslin, and Newport (1996) exposed 8 month old babies to 2 min of continuous speech made up of four three-syllable nonsense words repeated in a random order. The infants demonstrated sensitivity to the serial structure regularities, as they listened significantly longer to new words (novel combinations of the previously presented syllables) than to the trained words in the subsequent test phase. With respect to temporal structure, Nazzi, Bertoncini, and Mehler (1998) exposed new-born infants to sentences that were low-pass filtered to preserve only the prosody, and the behavioral measure was the rate of sucking on a pacifier. After habituation to sentences in one rhythm class, significant sensitivity to temporal structure was observed as an increase in sucking rate when sentences in a different rhythm class were presented in the test phase. Finally, for abstract structure, Marcus, Vijayan, Bandi Rao, and Vishton (1999) exposed 7-month old infants to a two-minute synthesized speech sample of 16 sentences that followed an ABA grammar such as “ga ti ga” and “la ni la.” In the transfer test phase, infants were exposed to speech made up of entirely new sentences that were either consistent with the ABA structure (such as “we fe we”) or inconsistent (“we fe fe”). Fifteen of 16 infants demonstrated a looking preference for inconsistent sentences, indicating that they had acquired and transferred knowledge of the abstract structure.

How can these regularities be learned? An important class of sequence learning models demonstrates the capability to predict future events by encoding the context or the history of previous events via recurrent- or self-connections (e.g., Cleeremans

& McClelland, 1991; Dominey, 1995; Dominey, Arbib, & Joseph, 1995; Elman, 1990). In these recurrent networks, the context is sensitive to events several positions in the past, allowing the networks to resolve ambiguities, as in determining the successor to B in the sequence ABCBAC. However, in many of these models, reliance on back-propagation learning in the recurrent connections places severe limitations on the networks' capacity to encode temporal structure. An additional shortcoming is that this recurrent context mechanism appears to be insufficient for representing the abstract structure common to the two isomorphic sequences ABCBAC and DEFEDF.

2. Cognitive sequence processing model

The temporal recurrent network (TRN) (Fig. 1) falls into the general category of recurrent networks described above (Dominey, 1995, 1998a; Dominey et al., 1995). Its development originated in an effort to reconcile behavioral and neurophysiological data from non-human primates in sequence learning tasks. Barone and Joseph (1987) observed that while monkeys perform a visuo-motor sequencing task, neurons in their prefrontal cortex (PFC) encoded both the spatial location and the relative sequential order of the spatial sequence targets. We postulated that visual inputs and recurrent cortico-cortical inputs to PFC (Goldman-Rakic, 1987) should be sufficient to yield this coding capability (Dominey et al., 1995). We also postulated that the coding of sequential structure in PFC would become bound to behavioral responses via plasticity in cortico-striatal synapses (i.e., synapses linking cortex to the input structure of the basal ganglia—the striatum), and that this plasticity would be modulated by the reward-related release of dopamine in the striatum (see Centonze, Picconi, Gubellini, Bernardi, & Calabresi, 2001 for up to

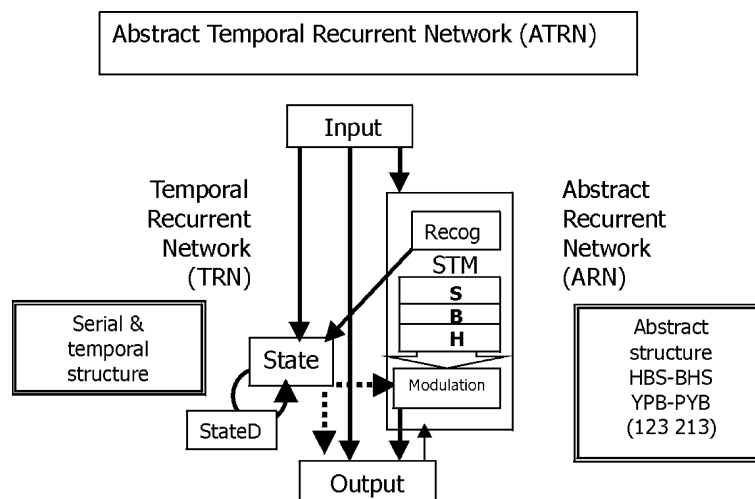


Fig. 1. Abstract temporal recurrent network (ATRN) model. The temporal recurrent network (TRN) exploits recurrent network dynamics to encode serial and temporal structure. To encode the abstract structure common to isomorphic sequences such as HBSBHS and YPB-PYB, the abstract recurrent network (ARN) stores the N previous elements of the current sequence in a short term memory (STM). The Recognition function compares the current sequence element response generated in Output to the previous elements coded in STM to detect the abstract repetitive structure. This abstracted coding is represented in the recurrent State. In the learned expression of abstract structure knowledge, the contents of STM are selectively provided to the output stream by the activation of Modulation neurons by State neurons.

date confirmation). The resulting recurrent network thus reproduced both the monkey's sequencing performance, and the neural activity of PFC neurons (Dominey et al., 1995).

In Fig. 1, PFC is represented by the recurrent network made up of State and StateD. Like the PFC of the primate, it is characterized by sensory inputs and recurrent cortico-cortical connections (Eqs. (1.1) and (1.2)). The striatum corresponds to Output in Fig. 1, and the modifiable cortico-striatal connections are thus the connections w^{SO} . Input, Output, State and StateD each consist of 5×5 arrays of leaky integrator units whose response latency (i.e., reaction time) is a function of the strength of their input signals. Eq. (1.1) describes the leaky integrator, $s()$, corresponding to the membrane potential or internal activation of State. In Eq. (1.2) the output activity level of State is generated as a sigmoid function, $f()$, of $s(t)$. The term t is the time, Δt is the simulation time step, τ is the leaky integrator time constant. As τ increases with respect to Δt , the charge and discharge times for the leaky integrator increase. An important characteristic of these leaky integrator neurons is that their activation time (or reaction time) depends on the strength of their inputs that can be modified by learning.

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

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

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—which correspond to glutamatergic and GABAergic cortical neurons and interneurons, respectively (Goldman-Rakic, 1987)—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 (Eqs. (2.1) and (2.2)).

$$\text{sd}_i(t + \Delta t) = \left(1 - \frac{\Delta t}{\tau}\right) \text{sd}_i(t) + \frac{\Delta t}{\tau} (\text{State}_i(t)), \quad (2.1)$$

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

Each element in a sequence is presented in Input, which thus generates a response in Output (Eqs. (3.1) and (3.2)). When a response is generated, active units in State encode the current sequence context, and connections from these State units to units in Output coding the response are strengthened (or weakened for incorrect responses), thus binding the sequence context to the current element in the sequence in an associative memory (Eq. (4)).

$$o_i(t + \Delta t) = \left(1 - \frac{\Delta t}{t}\right) o_i(t) + \frac{\Delta t}{t} \left(\text{Input}_i(t) + \sum_{j=1}^n w_{ij}^{SO} \text{State}_j(t) \right), \quad (3.1)$$

$$\text{Out} = f'(o(t)), \quad (3.2)$$

$$w_{ij}^{SO}(t + \Delta t) = w_{ij}^{SO}(t) + R * \text{State}_i * \text{Out}_j. \quad (4)$$

From a neurophysiological perspective, this associative memory is based on plasticity in the synapses linking cortical outputs to the striatum of the basal ganglia (Alexander, DeLong, & Strick, 1986). It has been demonstrated that when behavioral rewards (or events that predict rewards) occur, dopamine is released in the striatum (Ljungberg, Apicella, & Schultz, 1992) modulating the synaptic potentiation of NMDA receptors in cortico-striatal synapses (Walsh & Dunnia, 1993, and see Centonze et al., 2001).

2.1. Learning and reaction times

The result of this learning is that the next time this same pattern of sequence context activity in State occurs (i.e., the next time the sequence of elements that precede the current element is presented) the strengthened State–Output connections will cause State to increasingly activate the appropriate Output unit, predicting the current sequence element. This increased input from State results in a reduced reaction time (RT) for this response in Output. By this mechanism, the learning of serial structure will be demonstrated as RT reductions for all elements in learned sequences, as revealed in serial reaction time (SRT) tasks (Dominey, 1998a, 1998b). An example of this type of learning in the domain of infant language acquisition is illustrated in Fig. 2A where we demonstrate that the TRN displays sensitivity to serial order (Dominey & Ramus, 2000) comparable to that of 8-month old infants in the study of Saffran et al. (1996).

2.2. Learning temporal structure

In the simulation of primate behavior, it was important to be able to exactly reproduce the behavioral paradigms used. This required that stimuli were presented for particular time durations, with particular delays between successive stimuli. For this reason, we chose to use fixed, randomly chosen connections in the recurrent State network, so that learning-related changes would not interfere with the

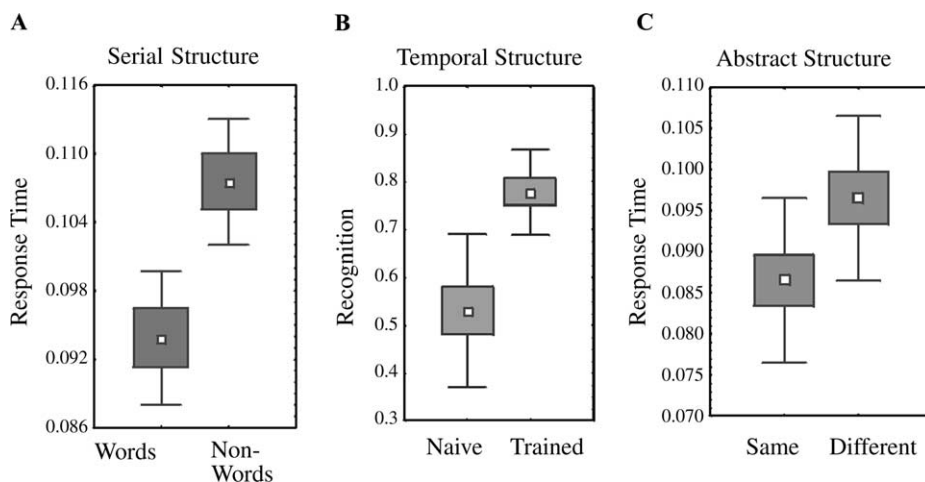


Fig. 2. Simulation of sensitivity to serial, temporal and abstract structure of sound sequences in the infant. (A) Simulation of Saffran et al. (1996). After exposure to the training sequence, RTs for words are significantly reduced compared to those for non-words made from the same syllables. (B) Simulation of Nazzi et al., 1998. Habituation with consonant–vowel sequences encoding prosodic structure of languages from different rhythm classes yields significantly greater discrimination between these languages than in the naive condition. (C) Simulation of Marcus et al., 1999. After exposure to the training sequences based on a given abstract structure, RTs for new sequences adhering to this abstract structure are significantly inferior to those for sequences made from the same elements but not adhering to the abstract structure.

temporal dynamics in the State network. The result was that we thus constructed a system that is sensitive to serial and temporal structure, and that can differentiate between two sequences that are identical in the serial order of the elements and differ only in their temporal or “rhythmic” structure (Dominey, 1988a, 1988b). An example of this type of learning in the domain of infant language acquisition is illustrated in Fig. 2B where we demonstrate that the TRN displays sensitivity to temporal structure that allows it to discriminate between languages from different rhythmic classes (Dominey & Ramus, 2000), comparable to performance of new born infants in the study of Nazzi et al. (1998).

2.3. Learning abstract structure

To permit the representation of abstract structure as we have defined it, the model must be capable of comparing the current response with previous responses to recognize repetitive structure (i.e., u, u, u, n-2, n-4, n-3 for ABCBAC, where “u” signifies unpredictable, and “n-2” indicates a repetition of the element 2 places behind, etc.). These functions would rely on the associative areas of the anterior cortex, and would permit the generalization of grammar-like rules to new, but “legal” sensorimotor sequences (Dominey, 1997; Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998).

To make this possible, we introduced a short term memory (STM) mechanism that is continuously updated to store the previous 7 ± 2 responses (Eqs. (5.1) and (5.2)), and a Recognition mechanism that compares the current response to the stored STM responses to detect any repeated elements (Eq. (6)). Working or short term memory and its use in recognition of previously presented stimuli in different modalities are characteristic functions of the anterior cortex as revealed in single unit recordings in the primate (Goldman-Rakic, 1987). Thus, the addition of an STM and associated recognition function remains clearly in the domain of the neurally plausible.

$$\text{for } i = 5 \text{ to } 2, \text{ STM}(i) = \text{STM}(i - 1), \quad (5.1)$$

$$\text{STM}(1) = \text{Out}, \quad (5.2)$$

$$\text{Recognition}_i = \text{STM}(i) * \text{Out}. \quad (6)$$

In Eq. (6), Recognition_i is non-zero if Out matches $\text{STM}(i)$. This permits the recoding of sequences in terms of their abstract structure that is now provided as input to State in the updated Eq. (1.1*).

$$\begin{aligned} s_i(t + \Delta t) = & \left(1 - \frac{\Delta t}{\tau}\right) s_i(t) + \frac{\Delta t}{\tau} \alpha \sum_{j=1}^n w_{ij}^{\text{IS}} \text{Input}_j(t) + \beta \sum_{j=1}^n w_{ij}^{\text{SS}} \text{StateD}_j(t) \\ & + \chi \sum_{j=1}^n w_{ij}^{\text{OS}} \text{Out}_j(t) + \delta \sum_{j=1}^n w_{ij}^{\text{RS}} \text{Rec}_j(t). \end{aligned} \quad (1.1*)$$

In this configuration, the factors α and χ will be set to zero, so that State is influenced only by Recognition and the recurrent connections. Now, in terms of the recoded abstract structure representation provided to State, the two sequences ABCBAC and DEFEDF are equivalent: u, u, u, n-2, n-4, n-3. For sequences that follow this “rule,” the pattern of activation produced in State by sub-sequence u, u, u, n-2 will reliably be followed by that context associated with n-4. To exploit this predictability, the system should then take the contents of the STM for the n-4th element and direct it to the output, yielding an RT reduction. This is achieved in the following manner: For each STM element (i.e., the structures that store the n-1, n-2, ... responses) there

is a unit that modulates and directs the contents of this structure to Output (Modulation units in Fig. 1). If one of these units is active, the contents of the corresponding STM structure is directed to Output as described in the last term of modified Eq. (3.1*).

$$o_j(t + \Delta t) = (1 - \Delta t/t) o_j(t) + \Delta t/t \left(\alpha \text{Input}_j(t) + \beta \sum_{j=1}^n w_{ij}^{\text{SO}} \text{State}_j(t) + \chi \sum_{k=1}^m \text{Modulation}_k \text{STM}(k)_i \right). \quad (3.1*)$$

Now, during learning, each time a match is detected between the current sequence element and an element stored in the STM, the connections between State units encoding the current context and the modulation unit for the matched STM element are strengthened as described in Eq. (7).

$$w_{ij}^{\text{SM}}(t + 1) = w_{ij}^{\text{SM}}(t) + \text{State}_i * \text{Recog}_j. \quad (7)$$

The result is that the next time this same pattern of activation in State occurs (e.g., before a match $n-4$ corresponding to the learned rule), the appropriate Modulation neurons will be activated (Eq. (8)) and the contents of the appropriate STM element will be directed to Output (Eq. (3.1*)) in anticipation of the predicted match, thus yielding a reduced reaction time. Thus, State encodes the abstract representation provided by Recognition, and provides access to the actual sequence elements stored in the STM via the influence of Modulation on Output.

$$\text{Modulation}_i = \sum_{j=1}^n w_{ij}^{\text{SM}} \text{State}_j(t). \quad (8)$$

An example of abstract structure learning in the domain of infant language acquisition is illustrated in Fig. 2C where we demonstrate that the ARN displays sensitivity to abstract structure comparable to that of 7-month old infants in the study of Marcus et al. (1999). Adult and simulation performance in serial and abstract SRT learning are examined in detail in Dominey et al. (1998).

2.4. Sensitivity to serial, temporal and abstract structure in the infant

At this point we have seen that the combined ATRN model can simulate infants' sensitivity to serial, temporal and abstract structure in the context of early language acquisition as revealed by the experiments of Saffran et al. (1996), Nazzi et al. (1998) and Marcus et al. (1999), respectively.

The ATRN model is based on neural implementation of a recurrent network, short term memory, recognition and modulation capabilities. There are no production rules, nor symbolic processing constructions. Rather, there are mechanisms that detect and encode structural regularity, allowing the system to adapt to structure to which it is exposed. Our next objective is to demonstrate how, through training, these learning capabilities can be used to form the basis for the acquisition of a syntactic comprehension capability.

3. Thematic role assignment (TRA) in syntactic comprehension

We must first clearly specify a target behavior that can be used to evaluate TRA. TRA can be functionally defined as the mapping of nouns onto their respective roles

Table 1
Nine sentence types as specified in Caplan et al. (1985)

Two place verb sentences
Active (A): The elephant hit the monkey.
*Passive (P): The elephant was hit by the monkey.
Cleft-Subject (CS): It was the elephant that hit the monkey.
*Cleft-Object (CO): It was the elephant that the monkey hit.
Three-place verb sentences
Dative (D): The elephant gave the monkey to the rabbit.
Dative pass (DP): The elephant was given to the monkey by the rabbit.
Sentences with two verbs
Conjoined (C): The elephant hit the monkey and hugged the rabbit.
*Subj-Obj Rel (SO): The elephant that the monkey hit hugged the rabbit.
Obj-Subj Rel (OS): The elephant hit the monkey that hugged the rabbit.

Non-canonical word ordering indicated by *.

in the thematic frame of the verb, corresponding to *agent*, *object*, *recipient*, etc., thus realizing the crucial comprehension function of determining “who did what to whom.” In a simplified manner, we can consider that in languages like French and English, there is a default or canonical order in which thematic roles are assigned (e.g., ‘Agent Object Recipient’ in English for the canonical sentence ‘John gave the ball to Mary.’) However, in non-canonical sentences (e.g., ‘The ball was given to Mary by John.’), this ordering is transformed, and thematic role assignment is guided, in part, by function items (e.g., closed class words including prepositions ‘to’ and ‘by,’ grammatical morphemes, etc.).

The ability to assign thematic roles can be quantified in clinical tests such as that developed by Caplan, Baker, and Dehaut (1985). The test consists of nine sentence types of varying syntactic complexity, five canonical and four non-canonical (see Table 1). Five sentences of each type are used for a total of 45 sentences. The sentences are constructed so that no semantic interpretation can contribute to the role assignment, which must proceed entirely as guided by syntactic function items. Sentences are read aloud to the patients in a pseudo-random order, and after each sentence, the patient should indicate by pointing at photographs ‘who did what to whom,’ indicating in canonical order the agent, object and recipient.

Using this material, Caplan et al. (1985) observed that in aphasic subjects (excluding severe “Wernicke’s” and “global” aphasics) canonical sentences of the type cleft subject (e.g., It was the elephant that hit the monkey) and active types (e.g., The elephant hit the monkey) were processed with the fewest errors, while non-canonical dative passive (see above) and subject–object relative types (The elephant that the monkey hit hugged the rabbit) sentences were processed with the most errors. We will employ this well-studied task to evaluate the TRA performance of the model. For a given input sentence, the model should respond by indicating the agent, object, and recipient in that canonical order.

4. Simulating thematic role assignment (TRA) in syntactic comprehension

A major hypothesis that drives our simulation approach is that sentences are a form of self-identifying data structure. That is, by relative combinations of word order, and explicit grammatical marking (by function words “by, to, from,” etc. or grammatical morphemes) a given sentence in any language contains the information necessary to perform correct TRA. Thus, an architecture that is sensitive to the

structure of word order and grammatical marking shall provide the basis for a language-independent TRA system. In this context, the transformation from the non-canonical “The cat¹ was introduced to the monkey² by the dog³” to the canonical “The dog³ introduced the cat¹ to the monkey²” as required for the Caplan task corresponds to the abstract structure 123-312, guided by a specific configuration of function words.

Based on this hypothesis, to realize the TRA task, a system should first be capable of distinguishing content words that will fill the thematic roles, from function words (or morphemes) that guide the syntactic analysis. This is likely one of the first acquisition processes in children, as by a process of “prosodic bootstrapping,” newborn infants can use acoustic cues to perform this lexical categorization (Shi, Werker, & Morgan, 1999). Across languages infants exploit acoustic, prosodic and distributional cues in continuous speech to perform this grammatical categorization (Morgan & Demuth, 1996). In this context we have demonstrated that the TRN learns to perform this distinction at greater than 75% accuracy, based on the temporal structure of phoneme duration and F0 variation in French (Blanc & Dominey, 2001; Blanc & Dominey, in preparation). Later, in the adult, function and content words are processed in a dissociated manner as revealed by numerous studies of psycholinguistics, aphasia and language related ERPs (e.g., Brown, Hagoort, & ter Keurs, 1999; Friederici, 1985; Neville, Mills, & Lawson, 1992; Osterhaut, 1997; Pulvermüller, 1995).

The system must also be able to store the content words in a working memory, and then to access this memory (potentially in an order different from the input order), guided by the function items, so as to re-establish the agent, object, recipient canonical ordering. This capability to access the content words in the canonical order, guided by the function words, provides the basis for the assignment of thematic roles. We can now allocate these functions to the ATRN model.

The desired TRA behavior for the model is to reproduce human-like performance in Caplan’s syntactic comprehension task. Sentences are presented, one word at a time to the model. Then, all of the nouns are presented together, and the model must respond by choosing first the agent, then the object, etc. in the canonical order, always choosing from the remaining unassigned nouns. Initially, the choice performance is random, and the model progressively learns to perform the syntactic comprehension task.

The nine sentence types from Caplan et al.’s (1985) syntactic comprehension task are coded so that each word corresponds to one element in the 25 element input array. Open and closed class items are processed by dissociable streams. This is “neurally” reasonable in that the TRN and other models (e.g., Shi, Morgan, & Allopenna, 1998) are capable of learning this distinction. Closed class words are directed towards the recurrent network (State) which thus maintains an ongoing representation of the syntactic structure of the sentence currently being processed. In Eq. (1.1*), parameters α , β and χ are non-zero indicating that State is influenced by the closed class words in the input stream, the recurrent connections from StateD, and Recognition. The open class words are stored in the STM in the order of their arrival. Thus for this task, Eq. (5.2) is modified so that STM is fed directly by the input rather than the output stream. Based on the syntactic context defined by the sequence of closed class words, the model must learn to successively choose the correct open class elements in their canonical order after the sentence is presented. Thus output is driven both by the Open class stream from the Input layer, and from the learning-driven modulation of STM contents, corresponding to non-zero values for parameters α and χ in Eq. (3.1*). In this manner, thematic role assignment is transformed into an abstract sequencing task.

4.1. Training

During the initial training phase, a given sentence is presented, one word at a time. Once the sentence is presented, the model is then provided with the open class nouns, one at a time, in their canonical order so that the correct associations can be learned. (In testing, the open class items are presented together, and the model must choose them in the correct order.) This training is functionally equivalent to providing, for each sentence, a scene depicting the nouns in their corresponding thematic roles. The model must learn these explicitly available syntactic-to-thematic mappings for the nine different sentence types. During this learning phase, the model learns to associate the different syntactic structures, in the form of distinct patterns of closed class words (encoded in the neural activity of the State layer), with the appropriate (re)ordering of open class words in the STM, by modifying connection strengths between State and Modulation neurons (Eq. (7)).

For an example sentence, “The elephant was shown to the monkey by the dog,” after the sentence is presented, the model is then provided with the first canonically ordered thematic role, i.e., the agent which in this case is ‘dog.’ This matches with the contents of the 3rd element of the STM. At this point, activity in State represents the sequence of function words that was provided. A learning signal thus strengthens connections between the current set of active neurons in State and the Modulation neuron that will modulate the contents of the third element of the STM into the output. The result of such learning is that when the same type of sentence is presented, the active State units will activate the appropriate Modulation unit, directing the contents of the 3rd STM element to the output, thus providing the agent response.

As the supervised training example proceeds, the next canonically ordered element, the object ‘elephant’ is provided. This matches with the first element of the STM, and the new State activity (which has been modified due to the match recognition) becomes associated with the modulation of the contents of the 1st STM element into the output, and so on for the recipient. This training is followed by a testing and correction period in which for each sentence the model must choose the correct agent, object and recipient from among the nouns that have not yet been assigned. Correct responses indicate that the learning has been effective and no connection strengths are changed. Errors result in weakening of the responsible State-to-Modulation connections, thus reducing the future probability of this error. The result of this training, for all nine sentence types, is that activity in the closed class stream (encoded in the evolving activity of State during sentence presentation) drives the selection of elements from the STM in the appropriate learned canonical order, allowing the model to perform the TRA task with 100% accuracy (Dominey, 2002).

5. Simulation conclusions and predictions

These simulation results are clearly oversimplified from a linguistic perspective, particularly regarding the status of the verb as an argument place holder. Still, the underlying representations are sufficient to allow the model to successfully perform the syntactic comprehension task. This suggests that cognitive sequencing processing may contribute to some of the processing involved in the transformation of syntactic structure to meaning in language comprehension.

This leads to two interesting and related observations. First, this indicates that at least for this set of nine sentence types, sufficient information is coded in word order and the configuration of function elements to allow a unique and unambiguous mapping of open class nouns onto their respective thematic roles. While this is not,

perhaps, extremely exciting, it is at least comforting to see an empirical demonstration that enough self referential information is coded in individual sentence types, and between sentence types, to permit a simple system to extract and learn the relevant regularities required for TRA.

The second observation is that, from the limited perspective of the model, language has the same processing status as that of any cognitive sequence system in which configurations of function elements govern the application of systematic transformations, as in domains including music, mathematics, logical theorem proving and artificial symbol manipulation tasks. If we consider that the model reflects some neurophysiological reality, then this observation thus leads to two rather daring predictions.

The first prediction is that the neurophysiological dysfunction that leads to impairments in syntactic comprehension should also lead to impairments in processing structurally analogous non-linguistic cognitive sequences that require the manipulation of complex abstract structure. Also, recall that the TRN can operate independent of the ARN (the ARN required adding new capability to the TRN). Thus, if abstract structure processing is specific to the ARN, then its impairment should not necessarily imply impairment in learning serial structure.

The second prediction is that in healthy subjects, brain imagery should reveal a common neurophysiological substrate for particular aspects of syntactic comprehension and for the corresponding non-linguistic cognitive sequence processing. The following sections address these predictions.

6. Prediction 1: Syntax and abstract structure processing are correlated in aphasia

We tested this prediction by studying the ability of 7 agrammatic aphasic subjects to process the abstract structure of non-linguistic letter sequences, and compared this to their syntactic comprehension performance using the Caplan sentence comprehension test (Dominey & Lelekov, 2000). In the non-linguistic abstract structure task, subjects were required to learn the abstract structure 123-213, by studying a set of 10 isomorphic sequences (letter strings) derived from this abstract structure. They were then asked to classify 20 new letter strings as corresponding, or not, to the learned target abstract structure. Subjects' performance in this abstract structure task was compared to their performance for syntactic comprehension, as evaluated by the nine sentence type "who did what to whom" task developed by Caplan et al. (1985). In support of our hypothesis, we observed (Fig. 3A) that agrammatic patients demonstrate performance impairments for syntactic comprehension and letter-sequence classification tasks that are significantly correlated ($r^2 = .86$, $p = .003$).

To verify that the patients' failure on the non-linguistic task was related to an impairment in processing complex or "non-canonical" forms that required transformation processing, we also tested these patients with the same abstract structure classification task, using the "canonical" abstract structure 123-123. We then compared canonical vs. non-canonical performance (in terms of percentage of correct responses) across the linguistic and non-linguistic tasks to test the prediction that processing of transformations in non-canonical order would be specifically impaired both for linguistic syntax and for non-linguistic abstract structures. For both tasks, non-canonical processing is selectively impaired with respect to canonical processing, as revealed by a significant effect for Order ($F(1, 5) = 31.7$, $p = .0025$), in the Task (linguistic vs. non-linguistic) \times Order (canonical vs. non-canonical) ANOVA. Most important, the Order \times Task interaction was not significant ($F(1, 5) = .053$, $p = .8$),

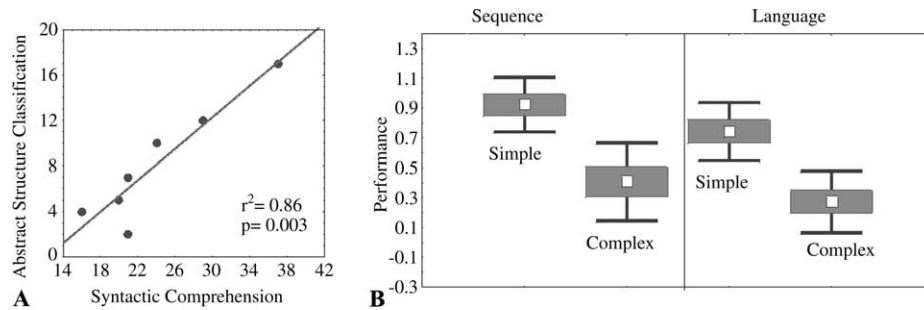


Fig. 3. (A) Correlation between processing impairments in treatment of complex abstract structure on the vertical axis, and syntactic structure on the horizontal axis for aphasic patients. (B) Breakdown of sequence and language processing by complexity for aphasic subjects. In both domains, there is a clear advantage for simple vs. complex forms.

indicating that this Order or transformation processing effect holds both for linguistic and non-linguistic tasks.

7. Prediction 1b: Aphasic patients learn serial but not abstract structure in an SRT task

In a related experiment, we compared the ability of aphasic patients to learn serial and abstract structure in a serial reaction time (SRT) task, with the prediction that they would succeed in learning serial structure (as recently demonstrated by Goschke, Friederici, Kotz, & van Kampen, 2001), but fail in learning abstract structure. We tested two right handed male aphasic patients with left perisylvian lesions, aged 50 and 52, and 5 healthy age matched control subjects. The two patients' syntactic comprehension impairment is revealed by their respective scores of 22/45 and 28/45 on the Caplan task (Lelekov-Boissard et al., 2000).

In the SRT protocol employed (Dominey et al., 1998), visual targets are serially presented on a touch sensitive screen, and reaction times are recorded for the successive touches. Classically, when such series of stimuli consist of repeating sequences, reaction times for the sequence elements are successively reduced (with respect to RTs for random series) as the sequence becomes increasingly well learned.

We know that implicit learning of serial structure of a repeating sequence yields uniform RT reduction for all sequence elements, and no transfer to new isomorphic sequences with the same abstract structure and different serial structure (Dominey et al., 1998). Learning abstract structure under explicit conditions yields additional RT reductions for abstract predictable elements (see Fig. 4), and transfer to isomorphic sequences (Dominey et al., 1998). Because optimal abstract structure learning requires explicit engagement in the task (Dominey et al., 1998; Gomez, 1997) subjects were tested in explicit conditions. Subjects' explicit knowledge of the target abstract structure was demonstrated by their matching the correct abstract structure to each of 6 isomorphic sequences reproduced from memory on the touch-screen.

Fig. 4 illustrates that agrammatic and control subjects learned the serial structure of the first sequence, as revealed by RT reductions for sequential (blocks 6 and 8) vs. random (block 7) stimuli. When exposed to a new isomorphic sequence in blocks 9 and 10, only the control group displays a significant transfer of performance for predictable elements, the signature of abstract structure learning and transfer (Dominey et al., 1998). Likewise, only the control group demonstrates a significant reduction of RTs for predictable vs. non-predictable elements.

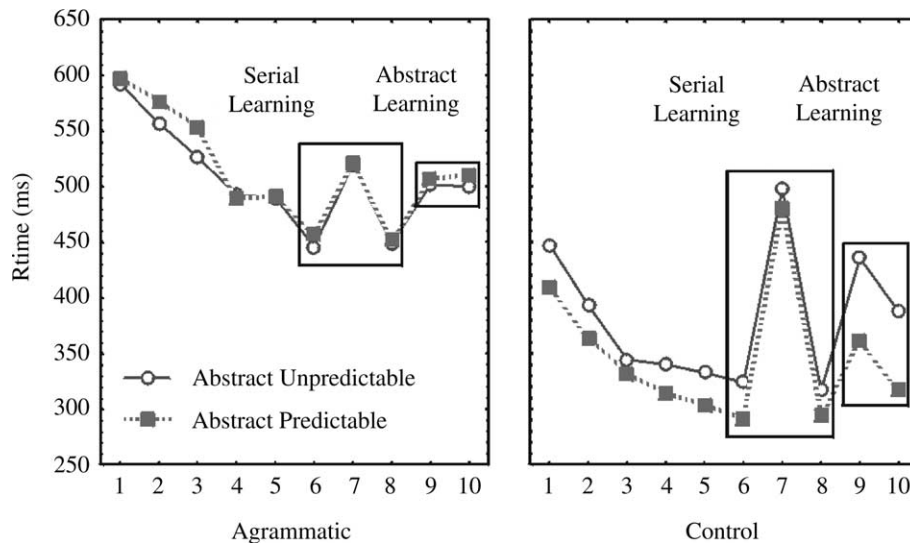


Fig. 4. Performance of agrammatic and control subjects on an SRT task that measures serial and abstract structure learning in 10 blocks of 108 trials (elements) per block. Blocks 1–6 and 8 use the repeating sequence ABCBACDEFEDF, that contains the abstract structure 123–213. Block 7 uses a random series to assay learning of serial structure. Blocks 9 and 10 use a new isomorphic sequence GBEBGEFDCDFC that shares the same abstract structure 123–213, but has a different serial structure. For such isomorphic sequences, elements corresponding to 123 are unpredictable by the abstract structure, while those corresponding to 213 are predictable, based on their fixed relations to elements 123. Reduced RTs for the unpredictable elements is due to learning of the serial structure. Additional reduction for predictable elements indicates learning of abstract structure that transfers to new isomorphic sequence blocks 9–10 for the Control but not Agrammatic patients.

These observations were confirmed by a multifactor ANOVA in which the independent variables were Group (Control, Agrammatic), Abstract Structure (Predictable, Unpredictable) and Block (1–10), and response time (RT) was the dependant variable. The Group \times Abstract Structure interaction was reliable ($F(1, 380) = 5.8, p < .05$), reflecting the observation of the advantage for elements that are predictable by the abstract structure only in the Control group. Planned comparisons between RTs for sequence block 6 and random block 7 revealed significant learning of serial structure both for control ($p < .001$) and aphasic ($p < .05$) groups. Planned comparisons ($p < .05$) confirmed the transfer of abstract structure to the new isomorphic sequence in blocks 9 and 10 only for the control group, with RTs for predictable elements in block 10 reduced to the same level as those for the learned sequence in block 8, and significantly reduced from those for random elements in block 7. Individual subject ANOVAs confirmed that serial structure learning occurred for all subjects, control and agrammatic, and that abstract structure learning and transfer occurred in all control subjects and in no agrammatic subjects. Thus, patients with left hemisphere lesions and associated impairments in syntactic comprehension are capable of learning serial structure in an SRT task, as recently demonstrated by Goschke et al. (2001), but they are impaired in extracting the abstract structure contained in the same sequences. These two experiments thus indicate that while the processing of serial order remains largely intact, the processing of systematic rule-based transformations as required for non-canonical sentences and abstract sequences are specifically impaired in agrammatic patients (see also Lelekov, Franck, Dominey, & Georgieff, 2000a).

8. Prediction 2: Syntax-related ERPs can be evoked in appropriate cognitive sequencing contexts

The “left anterior negativity” (LAN) is a language related ERP effect that has been revealed in response to particular forms of ungrammaticality, and in the processing of syntactic complexity associated with working memory costs (King & Kutas, 1995; Kluender & Kutas, 1993). This LAN effect has also been linked to the processing of closed vs. open class words (Brown et al., 1999; Neville et al., 1992; Osterhaut, 1997), particularly those associated with structural transformation processing (Rösler, Pechmann, Streb, Röder, & Hennighausen, 1998). Because of their syntactic function, closed class words are often associated with working memory loads in thematic role assignment (Kluender & Kutas, 1993). In this context we set out to generate a LAN response to function-symbols in abstract sequences (Hoen & Dominey, 2000). We thus developed a cognitive sequence processing task in which letter sequences are to be judged as acceptable or unacceptable based on pre-learned rules associated with special function symbols (analogous to function words) within the sequences (Hoen & Dominey, 2000).

The sequences were composed of 8 letters, successively presented for 400 ms followed by a 400 ms pause, yielding an 800 ms interval between the onset of two successive stimuli. In each sequence, elements 1–3 were three different letters (randomly chosen between A and V). Element 4 was a target element (either X or Z), and was the stimulus for which ERP analyses were performed. Elements 5–7 consisted of a second triplet of three different letters. Finally, element 8 was a repetition of the target element in position 4. Subjects were required to categorize the sequences as acceptable or unacceptable with respect to two conditions:

In the control condition, after presentation of elements 1–3 the letter Z appeared in target position 4. Subjects were asked to consider the sequence as acceptable if the position 8 element was a repetition of the target element Z in position 4. (Acceptable: ABCZDEFZ/Unacceptable: ABCZDEFG.) In the function-symbol condition, after presentation of elements 1–3 the letter X appeared in target position 4. Subjects were asked to consider the sequence as acceptable if the elements in positions 5–7 were a transformed version of elements 1–3 following the rule $123 \rightarrow 312$. (Acceptable: ABCXCABX/Unacceptable: ABCXBACX.) This condition fulfills the requirement that a specific symbol (X) allows the prediction of subsequent sequence organization based on a transformation of the previous context.

Fig. 5 illustrates the time course of activity at a representative left anterior electrode in response to the critical element in position 4, and the spatial topography of the Function-symbol minus control scalp voltage differences at 500 ms poststimulus. This subtraction reveals the significant lateralization of the negative shift observed in the 400–600 ms period for the Function-symbol vs. Control condition, with its central and left anterior predominance, similar to that of the syntactic LAN.

In the 0–200 ms post-stimulus period, the ERP profiles for the two conditions display a standard N1–P2 complex. Then, during the 200–400 ms window, for the Function-symbol condition, the P2 component remains positive until the onset of a negative shift at around 400 ms. In the control condition during this 200–400 ms period there is a negative going component followed by the onset of a P300-like positivity. During the 400–600 ms period, activity for the Function-symbol condition is characterized by a negative shift that appears more pronounced for the Left anterior vs. Right anterior electrodes. In contrast, during this 400–600 ms period, the control condition displays a symmetrical centroparietal positivity corresponding to the P300 effect. This indicates that the left asymmetry in the subtraction image (Fig. 5B) is due to the asymmetric negativity in the Function-symbol condition (Hoen & Dominey, 2000).

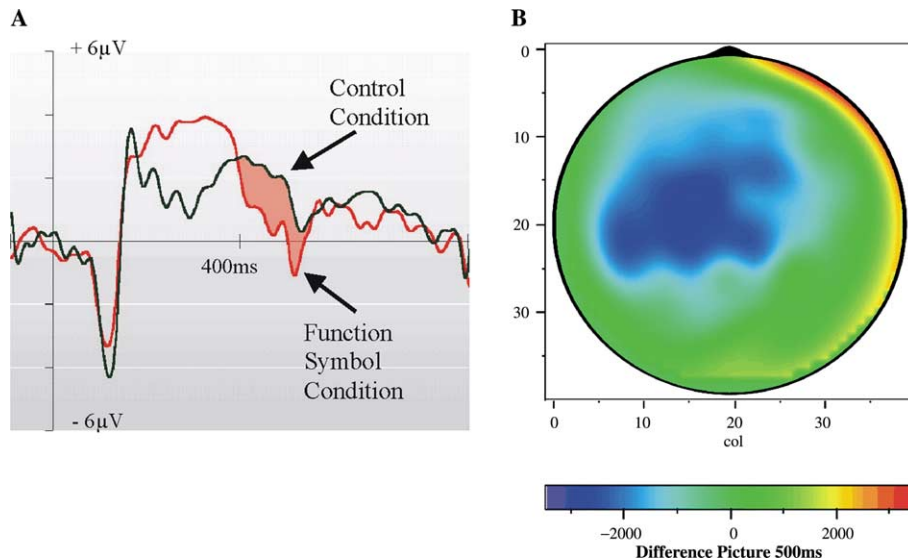


Fig. 5. (A) Grand averages for 10 subjects in the Function-symbol vs. Control Conditions for a representative Left Anterior electrode. Voltage scale from +6 to $-6 \mu\text{V}$. Time scale in milliseconds. Transformation condition shows a significantly negative going effect in the 400–600 ms time window with respect to control condition. (B) Subtraction image of Function-symbol minus Control conditions at 500 ms poststimulus. Note the left anterior distribution of the processing negativity (from Hoen & Dominey, 2000).

In the Function-symbol condition, the special symbol in position 4 provided information about the subsequent structure of the sequence, similar to the role of a grammatical function word. Likewise, this condition was also associated with a working memory load, in that the indicated transformation required storage and manipulation of the previous three elements. In the same sense, while function words define a specific lexical category, they are also quite often associated with syntactic processing that will require working memory (King & Kutas, 1995; Kluender & Kutas, 1993). Thus, the LAN effect we see is likely related both to this working memory processing effect and the “lexical” category of the function element.

While this lateralized negative-going potential is similar for function words and non-linguistic function symbols, it is noteworthy that we can observe differences between ERP responses to function symbols vs. those for function words in natural language. In particular, the function symbol is associated with a prolongation of the P2 component that remains positive until approximately 400 ms when the negative shift begins. A similar effect has been observed in an artificial grammar learning task (Baldwin & Kutas, 1997) where the processing of non-linguistic sequence elements that were consistent with the learned grammar elicited an enhanced positive going effect between 200 and 500 ms effect, quite similar to the effect seen in our experiment. This effect was considered to reflect the confirmation of a learned sequential expectancy. From this perspective, in our task, as contrasted with language processing tasks, subjects will have enhanced expectation that the 4th element will be either the function symbol, or the control cue, with attentional emphasis placed on the function symbol as it implies subsequent processing complexity. Thus, confirmed expectation of the function symbol in the context of the learned transformation would yield the extended P2 effect, as observed for grammatical targets in Baldwin and Kutas’s sequencing task.

9. Discussion

The forward linear progression of time imposes a fundamental sequential structure on all behavior, and thus the capacity to store and manipulate sequential information is of central importance for adaptive systems. Language is perhaps the most highly developed form of cognitive sequence processing and, as such, likely places the greatest demands on this capability. It appears quite clear that there exist neurophysiological processes that are specific and dedicated to language. The working hypothesis of this research is that, however, there is also a substantial functional and neurophysiological overlap between aspects of language processing and non-linguistic sequential cognition.

It is important to stress that we are not making the Piagetian claim that the organization of sensorimotor sequencing within the individual lays the basis for the sequential aspects of language processing (see Piatelli-Palmarini, 1994). Rather, we suggest, like many others (e.g., Greenfield, 1991; Lieberman, 1991; Rizzolatti & Arbib, 1998) that the development of neurophysiological processes for the control of complex sequential motor behavior has made a substantial functional contribution to the human linguistic capacity. In addition, we provide simulation, neuropsychological and neurophysiological evidence in support of this position. Similar approaches have recently studied music as a structure-containing analog of language with interesting results (Maess, Koelsch, Gunter, & Friederici, 2001; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). The power of our approach is that it introduces a new paradigm for the investigation of language processing based on cognitive sequencing that maintains the structural processing component of language, but has no “contamination” nor overlap with any existing cognitive system (such as music).

A cornerstone of this hypothesis was first laid out in a series of neural network simulation studies. Dominey et al. (1995) proposed a neural network model of recurrent connections in prefrontal cortex, and modifiable synapses linking cortex to basal ganglia, to explain electrophysiological activity and sensorimotor sequence learning behavior observed in the macaque monkey (Barone & Joseph, 1989). We subsequently explored the capabilities of this model in a number of cognitive sequencing tasks (Dominey, 1995; Dominey, 1998a, 1998b; Dominey & Boussaoud, 1997), and discovered that while it simulated with a remarkable fidelity human and non-human primate performance in serial and temporal structure learning, it failed to learn abstract sequential structure.

The abstract structure task was designed to test the ability to learn a rule that involved a systematic transformation and to apply this rule to new, isomorphic sequences. It was not initially conceived from a linguistic perspective. The addition of a short term working memory and comparison process resolved the abstract structure learning problem and lead to the ATRN cognitive sequence processing model described above (Dominey, 1997; Dominey & Ramus, 2000; Dominey et al., 1998).

The demonstration that the ATRN model could simulate human infants' sensitivity to serial, temporal (prosodic), and abstract structure of language or language-like sound sequences represented the first concrete evidence for the functional link between sequential cognition and language (Dominey & Ramus, 2000). At this point we realized that while the ARN had only been trained on one rule at a time, its transformation processing capacity could be extended to manipulate multiple syntactic transformations (i.e., sentence types), when guided by functional projections encoded in the TRN. We thus demonstrated that by allocating closed and open class processing to the TRN and ARN, respectively, the model could learn to associate different syntactic sentence types to the corresponding thematic role assignments

(Dominey, 2002). That is, cognitive sequencing laid the foundation for performing thematic role assignment. This completed the functional neurophysiological progression from sensorimotor sequence learning, via infant sensitivity to sequential structure, to adult syntactic comprehension.

To determine to what extent the model reflects the functional neurophysiology of human cognitive sequence and language processing, two strong predictions were proposed and tested. The first prediction was that dysfunction of the underlying neurophysiology associated with impairments in syntactic comprehension should also lead to impairments in the analogous non-linguistic cognitive sequencing tasks in aphasic subjects. Related to this was the idea that this deficit would be expressed for abstract structure but not for serial structure. The second prediction was that in healthy subjects, brain imagery (ERP) should reveal a common neurophysiological substrate for particular aspects of syntactic comprehension and of non-linguistic abstract cognitive sequence processing. In support of the first prediction, aphasic patients displayed strong correlations in their deficits in syntactic and abstract structure processing, while serial structure learning remained intact. With respect to the second prediction, ERP studies in healthy subjects revealed a left anterior negativity evoked by “function symbols” in an abstract structure processing task, analogous to that evoked by function words in language processing. Similarly, we have generated “abstract” P600-like responses to structure violations in abstract sequences (Lelekov-Boissard & Dominey, 2002; Lelekov, Dominey, & Garcia-Larrea, 2000b).

At this point, we must objectively introduce some strong reservations. It is likely disturbing for some readers to see that syntactic comprehension has been reduced to what is described in the simulations reviewed above. Admittedly, this represents a first approach towards simulation of neural processes involved in a limited version of thematic role assignment. Future testing and development of the model must include the treatment of a richer representation, particularly for the verb structure (e.g., Vosse & Kempen, 2000), a more detailed analysis of the influence of temporal structure on syntactic processing (Haarman & Kolk, 1991), and more realistic account of the temporal stages of syntactic processing (Friederici, 1995).

Despite its limitations, the model does have some merit. In particular it illustrates that a neurophysiologically grounded sequencing model can effectively perform accurate thematic role assignment on a limited set of sentence types based wholly on learning. Without modification the model should be able to learn the thematic role assignment task for multiple languages that rely differently on grammatical marking or morphology and word order. In the current case, simulation of the Caplan task revealed that based on this learning, the ATRN model could extract information from the relative configuration of function words to establish the associated mapping of nouns onto their respective thematic roles for nine different sentence types. The claim is thus, that to some extent, the model reflects at least some aspects of how this task is performed in humans, exploiting existing neural mechanisms for cognitive sequence processing.

Acknowledgments

This work was supported in part by funding from the ACI “Cognitique” of the French Minister of Research. T.L.B. and M.H. received support from the French Minister of Research, and J.M.B. received support from the Rhône-Alps Region “Emergence” Project. The authors acknowledge insightful comments from David Caplan, Peter Hagoort, H.H.J. Kolk, Laurie A. Stowe and an anonymous reviewer.

References

- Alexander, G. E., DeLong, M., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual review of Neuroscience*, 9, 357–381.
- Baldwin, K. D., & Kutas, M. (1997). An ERP analysis of implicit structured sequence learning. *Psychophysiology*, 34, 74–86.
- Barone, P., & Joseph, J.-P. (1989). Prefrontal cortex and spatial sequencing in the macaque monkey. *Experimental Brain Research*, 78, 447–464.
- Blanc, J. -M., & Dominey, P. F. (2001). *La prosodie comme lien entre discrimination lexicale et troubles du langage: Modélisation par un réseau récurrent*. Paper presented at the meeting: Journées Prosodie 2001, Grenoble France, 10–11 October 2001.
- Blanc, J. -M., & Dominey, P. F. (in preparation). *Prosody as a link between lexical discrimination and language disorders: Simulation by a Temporal Recurrent Network*.
- Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience*, 11(3), 261–281.
- 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. *European Journal of Neuroscience*, 13(6), 1071–1077.
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120, 235–253.
- Dominey, P. F. (1995). Complex sensory-motor sequence learning based on recurrent state-representation and reinforcement learning. *Biological Cybernetics*, 73, 265–274.
- Dominey, P. F. (1997). An anatomically structured sensory-motor sequence learning system displays some general linguistic capacities. *Brain and Language*, 59, 50–75.
- Dominey, P. F. (1998a). A shared system for learning serial and temporal structure of sensory-motor sequences? Evidence from simulation and human experiments. *Cognitive Brain Research*, 6, 163–172.
- Dominey, P. F. (1998b). Influences of temporal organization on transfer in sequence learning: Comments on Stadler (1995) and Curran and Keele (1993). *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24(1), 234–248.
- Dominey, P. F. (2002). A model of learning syntactic comprehension for natural and artificial grammars. In: E. Witruk, A.D. Friederici, & T. Lachmann (Eds.), *Basic functions of language reading and reading disability*. Dordrecht: Kluwer Academic Publishers, pp. 61–78.
- Dominey, P. F., Arbib, M. A., & Joseph, J. P. (1995). A model of cortico-striatal plasticity for learning oculomotor associations and sequences. *Journal of Cognitive Neuroscience*, 7(3), 311–336.
- Dominey, P. F., & Boussaoud, D. (1997). Encoding behavioral context in recurrent networks of the frontostriatal system: A simulation study. *Cognitive Brain Research*, 6, 53–65.
- Dominey, P. F., & Lelekov, T. (2000). Non-linguistic transformation processing in agrammatic aphasia. Comment on Grodzinsky: The neurology of syntax: Language use with and without Broca's area. *Behavioral and Brain Sciences*, 23(1), 34.
- Dominey, P. F., Lelekov, T., Ventre-Dominey, J., & Jeannerod, M. (1998). Dissociable processes for learning the surface and abstract structure of sensorimotor sequences. *Journal of Cognitive Neuroscience*, 10(6), 734–751.
- Dominey, P. F., & 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. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Friederici, A. D. (1985). Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition*, 19, 133–166.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain Language*, 50(3), 259–281.
- Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In Veron B. Mountcastle (Ed.), *Handbook of physiology* (Vol. 5, pp. 373–417).
- Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, 33, 154–207.
- Goschke, T., Friederici, A. D., Kotz, S. A., & van Kampen, A. (2001). Procedural learning in Broca's aphasia: Dissociation between the implicit acquisition of spatio-motor and phoneme sequences. *Journal of Cognitive Neuroscience*, 13(3), 370–388.
- Greenfield, P. M. (1991). Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, 14, 531–595.

- Haarman, H. J., & Kolk, H. H. J. (1991). A computer model of the temporal course of agrammatic sentence understanding: The effects of variations in severity and sentence complexity. *Cognitive Science*, 15, 49–87.
- Hoen, M., & Dominey, P. F. (2000). ERP analysis of cognitive sequencing: A left anterior negativity related to structural transformation processing. *NeuroReport*, 11(14), 3187–3191.
- King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376–395.
- Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbound dependencies. *Journal of Cognitive Neuroscience*, 5(2), 196–214.
- Lelekov, T., Franck, N., Dominey, P. F., & Georgieff, N. (2000a). Cognitive sequence processing and syntactic comprehension in schizophrenia. *NeuroReport*, 11(10), 2145–2149.
- Lelekov, T., Dominey, P. F., & Garcia-Larrea, L. (2000b). Dissociable ERP profiles for processing rules vs. instances in a cognitive sequencing task. *NeuroReport*, 11(5), 1129–1132.
- Lelekov-Boissard, T., Dominey, P. F., Chosson-Tiraboschi, C., Ventre-Dominey, J., Labourel, D., & Michel, F., et al. (2000). Selective deficits in processing syntactic and abstract sequential structure in agrammatic comprehension. *European Journal of Neuroscience*, 12(Suppl. 11), 168.
- Lelekov-Boissard, T., & Dominey, P. F. (2002). Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure. *Neurophysiol. Clin.*, 32, 72–84.
- Lieberman, P. (1991). Speech and brain evolution. Commentary on Greenfield. *Behavioral and Brain Sciences*, 14, 566–568.
- Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology*, 67(1), 145–163.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nat. Neurosci.*, 4(5), 540–545.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77–80.
- Morgan, J. L., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Erlbaum.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 1–11.
- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, 2, 244–258.
- Osterhaut, L. (1997). On the brain response to syntactic anomalies: Manipulation of word position and word class reveal individual differences. *Brain and Language*, 59, 494–522.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6), 717–733.
- Piatelli-Palmarini, M. (1994). Ever since language and learning: Afterthoughts on the Piaget–Chomsky debate. *Cognition*, 50, 315–346.
- Pulvermüller, F. (1995). Agrammatism: Behavioral description and neurobiological explanation. *Journal of Cognitive Neuroscience*, 7(2), 165–181.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, 21(5), 188–194.
- Rösler, F., Pechmann, T., Streb, J., Röder, B., & Hennighausen, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language*, 38, 150–176.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, 25, 169–201.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2), B11–B21.
- Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, 72, 105–143.
- Walsh, J. P., & Dunnia, R. (1993). Synaptic activation of N-methyl-D-aspartate receptors induces short-term potentiation at excitatory synapses in the striatum of the rat. *Neuroscience*, 57(2), 241–248.