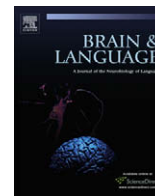




Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Neural network processing of natural language: II. Towards a unified model of corticostriatal function in learning sentence comprehension and non-linguistic sequencing

Peter Ford Dominey^{a,*}, Toshio Inui^{b,c}, Michel Hoen^a

^aINSERM U846 Stem Cell and Brain Research Institute, 18 ave. Doyen Lepine, 69675 Bron Cedex, France

^bDepartment of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan

^cERATO Asada Project, Japan Science and Technology Agency, Japan

ARTICLE INFO

Article history:

Accepted 11 August 2008

Available online xxxx

Keywords:

Syntactic comprehension

Non-linguistic sequence

Basal ganglia

Cortico-striatal system

Neural network

Language acquisition

ABSTRACT

A central issue in cognitive neuroscience today concerns how distributed neural networks in the brain that are used in language learning and processing can be involved in non-linguistic cognitive sequence learning. This issue is informed by a wealth of functional neurophysiology studies of sentence comprehension, along with a number of recent studies that examined the brain processes involved in learning non-linguistic sequences, or artificial grammar learning (AGL). The current research attempts to reconcile these data with several current neurophysiologically based models of sentence processing, through the specification of a neural network model whose architecture is constrained by the known cortico-striato-thalamo-cortical (CSTC) neuroanatomy of the human language system. The challenge is to develop simulation models that take into account constraints both from neuroanatomical connectivity, and from functional imaging data, and that can actually learn and perform the same kind of language and artificial syntax tasks. In our proposed model, structural cues encoded in a recurrent cortical network in BA47 activate a CSTC circuit to modulate the flow of lexical semantic information from BA45 to an integrated representation of meaning at the sentence level in BA44/6. During language acquisition, corticostriatal plasticity is employed to allow closed class structure to drive thematic role assignment. From the AGL perspective, repetitive internal structure in the AGL strings is encoded in BA47, and activates the CSTC circuit to predict the next element in the sequence. Simulation results from Caplan's [Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, 21, 117–175] test of syntactic comprehension, and from Gomez and Schvaneveldts' [Gomez, R. L., & Schvaneveldt, R. W. (1994). What is learned from artificial grammars?. Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 396–410] artificial grammar learning experiments are presented. These results are discussed in the context of a brain architecture for learning grammatical structure for multiple natural languages, and non-linguistic sequences.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

One of the astounding features of the human language acquisition capability is that it is pre-equipped to learn any human language. Over the years, developments in cognitive neuroscience, notably in brain imaging and lesion studies, have made it possible to examine the distributed neural networks underlying the language capability. Likewise, in an effort to isolate specific aspects of language-related processing, researchers have generated artificial language stimuli that preserve (or violate) properties that are hypothesized to be crucial to language. Together these types of

studies make a significant contribution to the understanding of the human language system. At the same time they pose a significant challenge to the computational neurosciences. That challenge is to develop simulation models that take into account constraints both from neuroanatomical connectivity, and from functional imaging data, and that can actually begin to perform the same kind of language and artificial syntax processing tasks.

1.1. The objective

Given these constraints on the functional neurophysiology of artificial and natural grammar processing, our task is now to map these constraints onto a neural network model that can perform these tasks. We have previously performed this type of exercise in the domain of the corticostriatal system for oculomotor

* Corresponding author. Fax: +33 4 72 91 34 61.

E-mail addresses: peter.dominey@inserm.fr (P.F. Dominey), inui@i.kyoto-u.ac.jp (T. Inui), michel.hoen@inserm.fr (M. Hoen).

saccades (Dominey & Arbib, 1992), and for learning sensorimotor associations and sequences in this system (Dominey, Arbib, & Joseph, 1995). We extended this sequence learning model to address a form of artificial grammar learning (Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998). In Dominey and Ramus (2000) we demonstrated how the resulting neural network model could simulate human-like behavior in the learning of serial, temporal and abstract structure of language as revealed by experiments with human infants (Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Nazzi, Bertoni, & Mehler, 1998; Saffran, Aslin, & Newport, 1996, respectively). We subsequently demonstrated how a model that combined the serial and abstract structure processing could learn to perform a thematic role assignment task with a variety of different grammatical constructions (Dominey, Hoen, Blanc, & Lelekov-Boissard, 2003), in three typologically distinct languages (Dominey, Hoen, & Inui, 2006).

The current research advances this effort in two fundamental ways: First, in contrast to the grammatical construction processing model of Dominey et al. (2006), the model presented here is based on a system of neural elements with membrane potential time constants, and thresholded firing rates. The result is that the system processes time (i.e., durations of inputs, reaction times, processing times) in a true analog manner, as opposed to a symbolic representation of time, which will be useful for word by word sentence processing. Second, in contrast to the language learning models developed in Dominey et al. (2003, 2006), where the functional neurophysiology of the system was handled at a relatively abstract level, in the current study we attempt to establish a much more detailed correspondence between processes of the model and their neurophysiological counterparts.

Finally, with the current model we address both syntactic comprehension, and the learning of artificial grammars in the classical sense of Reber (1967). In particular, the model is used to provide insight into behavioral results from an artificial grammar learning task described in Gomez and Schvaneveldt (1994). We will suggest that a common underlying neural system can provide a shared basis for the learning of natural and artificial grammatical structure. This will provide insight into current discussions of how different levels of sequential and grammatical structure are processed in the frontal cortex.

1.2. Characterizing thematic role assignment

In order to begin this modeling approach, we first characterize the behavior in question. A central function in syntactic analysis or syntactic comprehension is the assignment of thematic roles to noun phrases in sentences. 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.). More generally, across languages, a restricted set of cues including word order, grammatical morphology (including free morphemes or closed class words), and prosody will be sufficient to specify how the mapping is to be performed (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982; MacWhinney, 1982).

The ability to assign thematic roles has been quantified in different clinical tests used to assess agrammatism in aphasic patients suffering from lesions of the left cortical hemisphere. A well known version developed by Caplan, Baker, and Dehaut (1985) consists of 9 sentence types of varying syntactic complexity, five canonical and four non-canonical, and will serve as our target problem (see

Table 1). Five sentences of each type are used for a total of 45 sentences. Sentences are read aloud to the patients in a pseudo-random order, and after each sentence, the subject should indicate by pointing at photographs 'who did what to whom,' indicating in canonical order the agent, object and recipient. Interestingly, a rather significant subgroup of these patients with left hemisphere lesions demonstrate a deficit in thematic role assignment that is highly selective for non-canonical sentences. We note that 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.

In order to realize such a task, a system should first be capable of distinguishing function words (or morphemes) from content words. Numerous behavioral and event-related brain potential studies indicate that indeed, adults process function and content words in a dissociated manner (e.g., Friederici, 1985; 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 in a non-standard order (i.e., different from the input order) guided by the function items. This capability to reorder the content words, guided by the function words provides the basis for the assignment of thematic roles. In the simulation study of syntactic comprehension presented below, we demonstrate that after training on a supervised version of the Caplan task in which the correct thematic role assignment is provided, the model can then perform the standard unsupervised task correctly, including the generalization to new sentences.

2. Corticostriatal processing in language and non-linguistic sequences

Over the last decades the cognitive neurosciences have provided vast data on the functional neurophysiology of language and non-linguistic sequence processing. This has allowed the emergence of models that synthesize these results into coherent frameworks. Friederici (2002) has proposed a detailed model of the temporal processing in sentence comprehension. Phase 1 involves lexical categorization in superior and media temporal gyrus (STG and MTG) and the transfer of this information to the inferior frontal gyrus LIFG, and is associated with an early left anterior negativity in the 150–200 ms timeframe ELAN. In Phase 2 semantic and morphosyntactic integration take place in BA45/47 and BA44/45, respectively, associated with the N400 and left anterior negativity LAN, respectively, in the 300–500 ms time frame. In Phase 3 a late syntactic integration takes place in IFG and BA44 and left frontal operculum and the basal ganglia (Friederici & Kotz, 2003), associated with the P600. This phase is associated with reanalysis and repair processing. Such repair processing will require semantic working memory, which likely is provided by BA45/47 (Friederici, 2002).

Table 1

Nine sentence types as specified in the syntactic comprehension task of 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 a complimentary model, Hagoort has proposed a cortical continuum of integration/unification in his Memory, Unification and Control (MUC) model (2005), again paying particular attention to the link between function and neurophysiology. The mental lexicon, which allows retrieval of meaning based on lexical entries, is implemented in the temporal cortex and retrieves for each word its meaning and its structural frame which specifies the possible structural environment of that lexical item. Unification requires a combination of distinct forms of working memory for integration in the time domain. In this context, semantic integration will thus imply BA 45/47, and syntactic integration imply BA 44/45.

We have recently proposed a model of grammatical construction processing in which grammatical structure cues (based on MacWhinney, 1982) are encoded in a recurrent cortical network, and via corticostriatal associative memory retrieve the grammatical construction. This allows the mapping of semantic elements in BA45 onto the sentence level representation of meaning in BA44 (Dominey et al., 2006).

Our goal is to develop a neurophysiological model that begins to account for real-time aspects of sentence processing as specified in models including those proposed by Friederici (2002), the functional processing of the memory, unification and control model of sentence processing developed by Hagoort (2005), and the grammatical construction processing model of Dominey et al. (2006). Part of the novelty of this research is that it will also attempt to account for recent neurophysiological results concerning the processing of non-linguistic sequential material including artificial grammar learning (AGL) within the same framework. Above all, we want to situate this model in the context of distributed networks of cortical, basal ganglia and thalamic interaction, providing a more detailed link between processing and the underlying neurophysiology.

2.1. Corticostriatal function in language

Fig. 1 provides a functional neurophysiological overview of the model. In order to characterize the operation of this distributed network, we first specify the 'goal' of the system in terms of sen-

tence comprehension. Given an input sentence, the system should perform the necessary operations (e.g., lexical categorization, thematic role assignment, etc.) in order to represent the meaning of the sentence. How is meaning represented? Tettamanti et al. (2005) have demonstrated that listening to action-related sentences results in activation of frontoparietal motor circuits, including BA44, pars opercularis of the left inferior frontal gyrus (LIFG). Pullvermüller and colleagues (Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005) have demonstrated this link between left hemisphere representations for language and action, establishing a topographic link between linguistic representations of action and their corresponding motor representations. For the present purposes, let us consider that during sentence comprehension LIFG BA44/6 recruits lexical information at the word level and unifies this into overall representations that span multi-word utterances (Hagoort, 2005), and can lead to activation of the related motor areas (Hauk et al., 2004; Pulvermüller et al., 2005; Tettamanti et al., 2005).

Here we consider the succession of events in time, within the space of the distributed brain network, for the processing of words in a sentence, leading to the generation of this sentence level meaning representation in LIFG. In Phase 1, a process of lexical categorization takes place as the word form is categorized. Based on Friederici (2002) this is hypothesized to take place in the temporal lobe with closed class information being processed in the STG, and open class information processed in the MTG. Sakai (2005) reviewed data indicating that BA45/47 is the putative area for selection and integration of semantic information at the sentence/discourse level. This suggests the possibility of a working memory for these elements in BA45/47 which could allow their subsequent binding into the sentence level representation of meaning in BA44/6.

Information related to closed class words will be directed from STG to BA 47 in the IFG. In this context Inui, Ogawa, and Ohba (2007), and Ogawa, Ohba, and Inui (2007) have observed that closed class elements in Japanese (particles) selectively activate BA47. Based on the layered structure of BA45/47 we can also safely assume the existence of recurrent cortico-cortical connections

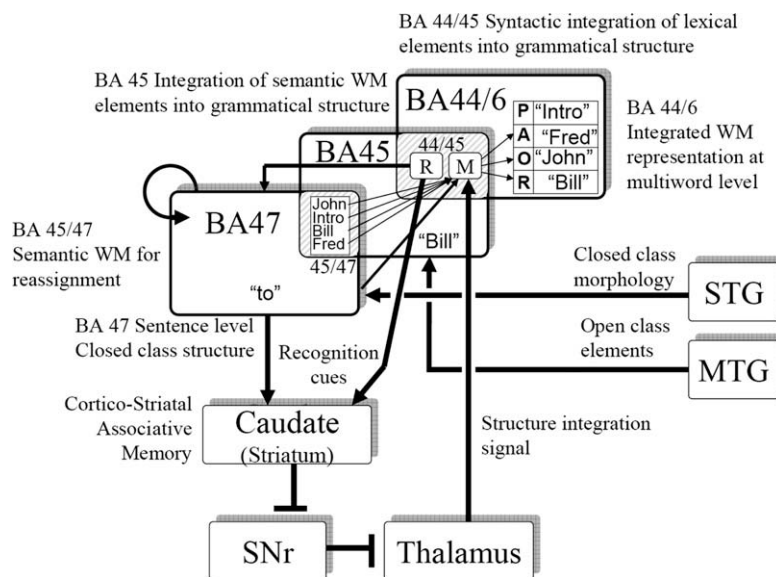


Fig. 1. Syntactic comprehension model architecture. Each of the labeled elements corresponds to a 5×5 array of leaky integrator neurons. Dual processing streams for closed and open class words, corresponding to STG and MTG, respectively. As the input sentence is read in 'John was introduced to Bill by Fred,' the recurrent network BA47 encodes the ordered history of closed class words from STG. BA45 receives open class elements from MTG. BA44/6 serves as a working memory encoding the meaning in terms of a 'predicate, agent, object, recipient' representation. Based on the syntactic context defined by the sequence of closed class words in the BA47-Caudate projections, the thalamic inputs to BA44/45M will modulate the open class element from BA45 to the correct population in BA44/6. In BA44/6 'Bill' is represented in the Recipient population, corresponding to the correct mapping for an input sentence 'John was introduced to Bill by Fred.'

within this area (Goldman-Rakic, 1987). These provide an important context preserving/encoding capability (Dominey et al., 1995), so that BA 47 activity represents the current closed class element within the overall context of the sentence. Continuing with the processing of closed class elements, we propose that corticostriatal projection neurons in BA 47 project to the head of the caudate, in a canonical cortico-striato-nigro-thalamo-cortical circuit. Recent data from human cortico-striatal projection tracing using diffusion tensor imaging (Lehéricy et al., 2004) and anterograde tracing (Wiesendanger, Clarke, Kraftsik, & Tardif, 2004) support the proposal that corticostriatal connections in humans are organized in multiple overlapping circuits, similar to that observed in non-human primates (Alexander, DeLong, & Strick, 1986). This leads us to propose a 'grammatical structure' circuit in the cortico-striatal system with distinct and segregated paths, similar to the proposal of Ullman (2006), addressed in more detail in Section 2.3.

2.2. Language learning in the cortico-striato-thalamo-cortical system

We hypothesize that via this circuit, the system can retrieve information related to the syntactic roles (or frames in the unification component of the MUC model of Hagoort) that are licensed by this word or group of closed class words. One potential mechanism for establishing this encoding, that will take place during language acquisition, is via modifiable cortico-striatal synapses that form the basis for a robust form of associative memory. While the broad discussion of poverty of the stimulus, etc. is beyond the scope of this paper (but see Dominey 2005a, 2005b), let us consider that at some given point during early acquisition (18–24 months of age), the child hears a sentence and observes an action-scene which that sentence accurately describes. Further consider that the child is familiar with the words in the sentence, but not with the particular grammatical form, e.g., a passive of the form 'John was introduced to Bill by Fred.' Observation of the action will lead to a representation of that action in a structured manner (encoding who did what to whom) within the LIFG BA44 (see Tettamanti et al., 2005).

When the subsequence 'to Bill' is processed, and the system has the pre-established representation of Bill as the Recipient of the action, then conditions are favorable for learning the relation between 'to' and assignment of the recipient role. The details of how this learning occur will be described in the next section. In summary, the encoding of the closed class word 'to' in cortex and the recognition of the thematic role played by 'Bill' generates activity in the caudate, which in turn inhibits SNr, disinhibiting thalamus and generating a final pattern of activation in BA 44/45. We hypothesize that via this activation, the thematic role for the current open class 'Bill' element is licensed to take the recipient of the verb 'introduced' (note that this is a language dependant operation—in other languages post-positions will license the word that precedes them). During learning, the system will initially operate by trial and error. Successful trials in which the (initially random) choice of the system is correct will result in strengthening of the responsible cortico-striatal connections. Likewise, incorrect choices (resulting in incorrect thematic role assignments) will result in weakening of the responsible cortico-striatal synapses. This dual mechanism (strengthening and weakening appropriately) thus favor repetition and consolidation of correct responses, and elimination of incorrect responses. It should be made clear that the functional neurophysiology of this form of learning is quite well documented in the domain of sensorimotor sequence learning. If we consider that correctly understanding a sentence will generate some internal reward, then we can hypothesize that dopamine release in the striatum during correct trials in learning favors synaptic strengthening via LTP (Calabresi, Picconi, Tozzi, &

Di Filippo, 2007). Likewise the absence of this expected reward in the face of misunderstanding will lead to reduction in the baseline activity of dopamine producing cells in the SNc and VTA (ventral tegmental area), and a corresponding weakening of these synapses (Calabresi et al., 2007). The net result of this BA 47-Caudate-SNC-Thalamus-BA44/45 projection is a modulation of the licensing of the relevant open class element encoded in BA45 for its corresponding thematic role, which will be encoded in BA44. Thus, this cortico-striatal loop participates in the syntactic unification by establishing the licensed thematic role assignment (TRA). This is consistent with observations of increased activity in BA44 and LIFG during tasks with high syntactic complexity (Caplan, Stanczak, & Waters, 2008; Moro et al., 2001; Stromswold, Caplan, Alpert, & Rauch, 1996). In conditions in which the assignments cannot be made as each open class word arrives, the system will require a form of syntactic working memory. Fiebach, Schlesewsky, Lohmann, von Cramon, and Friederici (2005) indeed observed activation of BA44 in conditions requiring such working memory. Given this 'long loop' CSTC pathway, we will consider that this reanalysis operation corresponds in part to the late Phase 3 integration of Friederici (2002) and Friederici and Kotz (2003).

2.3. Evidence for language-related cortico-striatal connectivity in human

Such a model would predict significant roles for striatum and thalamus in sentence processing, and from our perspective, they will be based on a 'grammatical structure loop' or set of parallel 'language loops' in the cortico-striatal system. Indeed, this prediction is validated in a number of brain imaging and neuropsychological studies. For example, Moro et al. (2001) observed significant activation of left BA45, and of the left caudate nucleus under conditions that specifically required syntactic processing. Similarly, during the processing of syntactic anomalies, Friederici and Kotz (2003) observed cortical activity in the left posterior frontal operculum adjacent to BA 44 and in the putamen of the left basal ganglia. Addition support comes from neurological studies. Insults to the caudate via lesion (Alexander, Naeser, & Palumbo, 1987; Pickett, Kuniholm, Protopapas, Friedman, & Lieberman, 1998), or Parkinson's and Huntington's disease result in a broad spectrum of language deficits including deficits in syntactic comprehension (Grossman et al., 1999; Hochstadt, Nakano, Lieberman, & Friedman, 2006; Teichmann et al., 2005) and grammatical rule processing (Ullman et al., 1997). Likewise, thalamic lesions can lead to linguistic deficits including deficits in sentence comprehension, and the use of syntactic complexity in production (De Witte, Wilsens, Engelborghs, De Deyn, & Mariën, 2006).

From the perspective of neuroanatomy, as stated above, Lehéricy et al. (2004) provide diffusion tensor imaging (DTI) data supporting a complex parallel cortico-striatal connectivity, though they did not specifically investigate Broca's areas. Leh, Ptito, Chakravarty, and Strafella (2007) provide further evidence from DTI studies in 6 humans revealing a massive topographic projection system from prefrontal cortex, inferior and middle temporal gyrus to the caudate. Again while specific connections with Broca's areas were not investigated, these regions fall within the large prefrontal regions that displayed cortico-striatal connectivity. Another DTI study examines thalamo-cortical connectivity and demonstrates that thalamic nuclei MD, VA and part of the anterior complex are connected to prefrontal cortex (Behrens et al., 2003). Postuma and Dagher (2006) performed a meta-analysis of coactivation of cortex and basal ganglia in order to reveal emergent patterns of functional connectivity. This study indicates that the left caudate is likely to be coactivated with dorsolateral prefrontal cortex and the inferior frontal gyrus. Di Martino et al. (2008) performed a resting state fMRI study to examine functional connectivity of the hu-

man striatum. Ventral caudate was associated with the IFG (BA47), and dorsal caudate with BA 45, 47. Finally, molecular studies of the FOXP2 gene have revealed that this gene is expressed in Broca's area and the striatum, and that both of these regions are structurally and/or functionally abnormal in members of the KE family who demonstrate a mutation in the FOXP2 gene (Vergha-Khadem, Gadian, Copp, & Mishkin, 2005). This indicates a functional connectivity between Broca's region and the caudate whose development is mediated by this gene.

2.4. Corticostriatal function in non-linguistic sequence and artificial language learning

In order to probe the neural substrates of language-related processing more precisely, one can develop stimuli that manipulate specific aspects of linguistic structure. Extrapolating this logic, one can use non-linguistic stimuli which possess structural characteristics that are hypothesized to play a role (or not) in language (see Petersson, Forkstam, & Ingvar, 2004). Thus, Friederici, Bahlmann, Heim, Schubotz, and Anwender (2006) exposed human subjects to letter strings that were generated by a phrase structure grammar and a finite state grammar. Their hypothesis was that the former embodies a principal characteristic of human language while the latter does not, and that this difference would be reflected in their selective recruitment in the left frontal cortex. Their fMRI results indicating that processing phrase structure grammar activates Broca's area (BA44) and the frontal operculum (FOP) while finite state grammar activates only the phylogenetically older frontal operculum. They subsequently performed DTI tract tracing from these two areas of interest indicating that the BA44 forms a network with the mid posterior portion of the superior temporal gyrus (STG) which is activated in processing syntactic complexity and the processing of the phrase structure grammar. In contrast, DTI tract tracing from the FOP area indicated connectivity with anterior portion of the STG, which together with the FOP has been reported for the processing of local phrase structure violations (Friederici et al., 2003). In a related study, Bahlmann, Gunter, and Friederici (2006) examined the same phrase structure (PSG) and finite state grammar (FSG) distinction in ERP studies. Among the FSG/PSG differences was a late positivity (characteristic of linguistic and certain non-linguistic grammatical structure processing) whose amplitude varied depending on the violation location only for the PSG.

In a related study using the classic artificial grammar paradigm of Reber (1967), Petersson et al. (2004) demonstrated activation of a left frontal network including BA's 6, 9, 44, 45 and 47 (see this paper for a review of this topic). In a subsequent study that involved more prolonged training, Forkstam, Hagoort, Fernandez, Ingvar, and Petersson (2006) revealed the presence of a corticostriatal network whose activity was correlated with learning and performance in an artificial grammar processing task. Left BA 45 activation was specifically related to syntactic violations and activation in the head of the caudate nucleus correlated positively with syntactic correctness after prolonged learning (8 days). Tettamanti et al. (2002) and Musso et al. (2003) generated test material from natural languages applied to non-native speakers, and demonstrated activation of BA44 only when these synthetic stimuli maintained syntactic vs. fixed position regularities.

In a pioneering study, Fletcher, Büchet, Josephs, Friston, and Dolan (1999) used an artificial grammar learning task which allowed dissociable study of item learning (within block) vs. rule learning (via explicit rule testing with feedback) across blocks. Over the course of the learning, they observed a shift from early item-based processing with right PFC activation to late rule and contextual based processing with increasing activation of left PFC. This left hemisphere PFC activation for abstract rule processing is consis-

tent with the hypothesis that this brain region might play a more general role in language and non-linguistic structural rule processing. Hoen, Pachot-Clouard, Segebarth, and Dominey (2006) tested subjects in grammaticality judgement of natural language sentences, and in non-linguistic sequences that required the use of a transformational rule. The areas of activation common to both tasks included a left frontal region adjacent to Broca's area (BA6/9/46). The rules employed in this task were of the type ABCxBAC where the x indicated the type of transformation that should be applied to the first triplet in order to generate the second, with the x thus serving the role of a non-linguistic function word. Given the distribution of cortical areas involved in these tasks from BA 47, 45, 44, 46, 6 and 9, we can consider the proposed integration gradient of Hagoort (2005), with different levels of structural processing distributed along this gradient.

In this context Koechlin and Jubault (2006) studied a complex sequencing task that involved embedded subsequence initiation and termination, guided by special symbols. They identified a functional hierarchy of processing where within-sequence processing was associated with bilateral activation of BA44, and the processing of signals linked to superordinate sequence initiation and termination were associated with bilateral activation of BA45.

With respect to second language representation, Wartenburger et al. (2003) tested semantic and grammatical processing in bilingual subjects with early vs. late acquisition of the second language. Their results indicate that for processing of grammar in the second language, late acquisition subjects display an increase in activation in bilateral IFG with respect to early acquisition subjects. Within the late acquisition group L2–L1 contrasts for grammatical processing revealed bilateral activation of the IFG and basal ganglia. This can be reconciled with the suggestion of Perani and Abutalebi (2005) in which a common neural system underlies acquisition of both, modulated by factors including age of acquisition and proficiency.

3. From neurophysiology to neural network model

We can now proceed with the specification of the simulation model. The model architecture is illustrated in Fig. 1. The cortical gradient of structural processing or unification is represented in the transition along BAs 47–45–44. Auditory or written word input undergoes an early lexical categorization, with closed class words making their way to BA47 via STG, and open class words to BA45 via MTG. BA45 encodes the current open class element, and a semantic working memory in BA45/47 maintains the sentence level ensemble of open class words available for possible reanalysis and repair. We model BA47 as a recurrent network which thus takes into account the current closed class element, as well as the previous closed class elements, with this history being encoded by information flow through the recurrent connections. BA47 projects to the head of the caudate, and via the 'canonical' cortico-striato-nigro-thalamo-cortico circuit, projects to BA44/45. While the current open class word is represented in BA45, the sentence level integrated meaning is represented in BA44/6 as proposed by Hagoort (2005). For simplicity, we assume a form of working memory with distinct neural populations corresponding to thematic roles, and distinct patterns of activity within those populations encoding the identity of the holder of that particular thematic role in the current sentence. This is represented in Fig. 1 with the Predicate, Agent, Object, Recipient roles indicated by PAOR. The relation between these two representations is critical: As part of syntactic integration, the word level meaning (BA45) must become integrated or bound—with the properly assigned thematic role—into the sentence level meaning in BA44/6. The assignment of a given open class element in a sentence to its thematic role will depend

on multiple factors including the particular grammatical construction, the local and global configuration of closed class elements in the sentence, the relative order of the open class element in the sentence. The system must take these factors into consideration.

Given these requirements, and the neuroimaging evidence for activation of BA44/45 during syntactic processing, we can consider that there is a relatively complex interaction between BA44 and BA45. We thus propose two functional subdivisions of this region. BA44/45 M (for modulation) modulates the transfer or binding of semantic content in BA45 into the grammatical/thematic frame of meaning at the sentence level in BA44/6. Thus, the thalamocortical input will modulate the connections from BA45 to BA44/6 such that the open class element coded in BA45 is transmitted to BA44/6 appropriately in the context of thematic role assignment (as for 'Bill' in Fig. 1). The second functional subdivision, BA44/45 R (for recognition) plays a different role—recognizing the open class element in BA45 and its role in the meaning coded in BA44/6 (as for 'Bill' in Fig. 1). These two areas thus play reciprocal roles: During learning BA44/45R detects the thematic role binding for the current open class element, (e.g., recognizing that 'Bill' in BA45 matches the Recipient (R) role in BA44/6), in the ongoing context of the sentence as encoded in BA47. BA44/45M provides the corresponding modulation or binding function, binding this element in BA45 to its proper role in BA44/6 during comprehension of the same type of sentence, when the same context is encoded in BA47. This recognition signal encoded in BA44/45R provides a crucial learning signal such that context encoding in BA47 becomes linked to the binding signal in BA44/45M (explained in the next paragraph). During comprehension, this activity in BA47 should drive the BA44/45M modulation population so that it modulates or binds the neural activity for the current open class element onto its corresponding thematic role—the same thematic role that was recognized in BA44/45R during learning in the same context as encoded by BA47.

Let us reconsider the situation of learning a new grammatical structure mapping, in which case the system has observed an event and encoded that event in BA44/6. The system is now learning how to map the sequential succession of open class elements (in the corresponding sentence) onto their thematic roles. BA44/45 R (for recognition) compares the current open class element in BA45 ('Bill') with the event representation of BA44/6, and detects a match with the R (Recipient) role. This information projects via a corticostriatal pathway into the caudate nucleus in a zone that overlaps with the projections from BA47. The BA44/45R input provides a teaching signal specifying that the current activity in BA47 should drive binding of BA45 contents to the Recipient role. Through corticostriatal plasticity, corticostriatal synapses on striatal neurons that are activated by both the BA47 inputs the BA44/45R inputs will become strengthened in an associative memory that binds the BA47 and BA44/45R inputs together. The result of the learning is that at some future time, when the same sentence type is being processed, the same input from BA47 will activate the caudate –SNr–thalamus–BA44/45M pathway and retrieve the associated binding such that BA44/45M binds or modulates the current open class element into the Recipient role. Importantly, this learning will generalize to new sentences that adhere to the learned grammatical structure. Via this mechanism, corticostriatal plasticity is employed to allow closed class structure to drive thematic role assignment.

4. Simulating language learning and syntactic comprehension

The model is constructed from ensembles of neuron-like units, each of which simulates the membrane potential of a neuron (or population of neurons) and generates an analog signal of firing/dis-

charge rate, as illustrated in Fig. 2. The model acquires the capacity to perform the syntactic comprehension task by learning, which results in the modification of connection strengths between different neurons. The desired behavior is to reproduce human-like performance in syntactic comprehension. Here we consider Caplan's task in which sentences are presented, one word at a time to the model which then responds by emitting the nouns in their canonical order when presented with the remaining unassigned nouns. In contrast to the online processing described above, this post hoc processing corresponds to a reanalysis of the sentence as when initial thematic role assignment has failed.

The 9 sentence types from Caplan et al.'s (1985) syntactic comprehension task are re-coded so that each word corresponds to one element in the 25 element input array of the model. The words are categorized into closed and open classes. Closed class words enter a processing stream directed towards the recurrent network (BA47) which thus maintains an ongoing context of the sequence of closed class words presented for the sentence currently being processed. The open class words enter a processing stream by which they are stored in a working memory for subsequent syntactic integration into the global representation in BA44/6.

The current simulation adheres to the Caplan et al. (1985) protocol, such that the input sentence is processed, and then the model is presented with the set of open class elements, and must choose first the agent, then the object and finally the recipient. This is realized in the model such that the semantic selection and integration will take contents from a semantic working memory in BA45/47, consistent with Sakai (2005).

Training: During the training phase, a given sentence is presented, one word at a time in a temporal sequence as just described. Once the sentence is presented, the model is then provided with the nouns/open class words, one at a time, in their canonical order. In Fig. 1, this corresponds to the Agent, Object and Recipient roles being successively filled in BA44/6. In the Caplan protocol, this would be equivalent to the experimenter herself pointing to the nouns in their canonical order after reading the sentence, i.e., showing the subject the correct responses. For canonical input sentences, the open class words are thus simply re-presented in the same order as in the initial input. For non-canonical input sentences, the nouns are transformed into their canonical order. During this learning phase, the model learns to associate the different patterns of closed class words (encoded in the neural activity of the BA47 layer), with the appropriate (re)ordering of open class words, through the modification of connections linking activity in BA47 with the thematic role binding neurons in BA44/45M. In the current model we propose that this interaction takes place via the corticostriatal pathway as shown. In Dominey et al. (2003) this was a direct connection.

For an example sentence, 'The elephant was transferred to the monkey by the dog,' after the sentence is presented, the contents of the semantic working memory BA45/47(1) is elephant, BA45/47(2) monkey, and BA45/47(3) dog. The model is then provided with the first canonically ordered thematic role in BA44/6, i.e., the agent which in this case is 'dog.' This matches with the contents of the BA45/47(3). A learning signal strengthens connections between the current set of active neurons in BA47 (whose activity is due to the sequence of function words that was provided) and a modulatory neuron BA44/45M(3) that will modulate the contents of the third element of the BA45/47, i.e., BA45/47(3), into the Agent role. The result of such learning is that when the same type of sentence is presented, the active BA47 units will activate the appropriate modulation unit, i.e., BA44/45M(3), selecting the contents of working memory BA45/47(3), 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 contents of BA45/47(1), and the new BA47 activ-

ity (which has been modified due to the match recognition) becomes associated with the activation of BA44/45M(1) that directs the contents of BA45/47(1) into the output representation in BA44/6, and so on for the recipient. The result of this training, for all 9 sentence types, is that activity in the closed class stream (encoded in the evolving activity of BA47 during sentence presentation) will drive Modulation units in an appropriate way that allows the selection of elements from the BA45/47 in the appropriate learned canonical order.

Testing and correction: After this type of training, we then proceed to a testing period. In the testing period, after the sentence has been presented, instead of providing the correct response for agent, object, and recipient, we force the model to make a choice between the remaining nouns that have not yet been assigned. For the sample sentence 'The elephant was transferred to the monkey by the dog,' after the sentence is presented, we then present the three open class elements 'elephant, monkey, dog' as potential agents, and the model must choose (i.e., activate the correct BA44/45M neuron), based on knowledge acquired in the training described above, the noun among these three that corresponds to the agent, which in this case is 'dog.' Correct responses indicate that learning has been effective, and no connection strengths are modified. Incorrect responses result in a weakening of the connections linking BA47 activity to BA44/45M activity, thus reducing the probability of the erroneous response being repeated. If the first response is correct, then the remaining two nouns (i.e., those that were not already chosen) are presented. The final choice is reduced to the only remaining noun that has not already been chosen. This corresponds to the Caplan task in which the subjects are asked to

point to the agent, object and recipient images, with a given image corresponding to only one of the roles. Details of the simulation of processing this sentence are illustrated in Fig. 2.

4.1. Simulation methods

A population of model 'subjects' was created by using the model architecture described in Fig. 1, and starting with different seed values of a random number generator used to assign synaptic weights to the recurrent connections. This population underwent a training procedure where each of the first 6 sentence types (two-argument and three-argument verb sentences—see Table 1) was presented approximately 10 times, in the training mode described above. That is, after the sentence was 'read' to the subject, the subject was then presented with the noun phrases, one after another, in their canonical order. After this training phase, subjects were exposed to a testing and correction period, also described above. During this period, after a sentence was read, the nouns that had not yet been assigned a thematic role were presented simultaneously and the model was forced to choose the correct one, with error-driven learning. During this testing period, each of the six sentence types was presented approximately 75 times.

Model subjects that completed this training and demonstrated no errors at the end of the training were then exposed to all nine sentence types in the same testing with error-driven learning, with approximately 60 exposures for each of the nine sentence types. This pretraining with a subset of the 9 sentence types follows the spirit of Elman (1993) in 'starting small,' and allows us to isolate potential good learners.

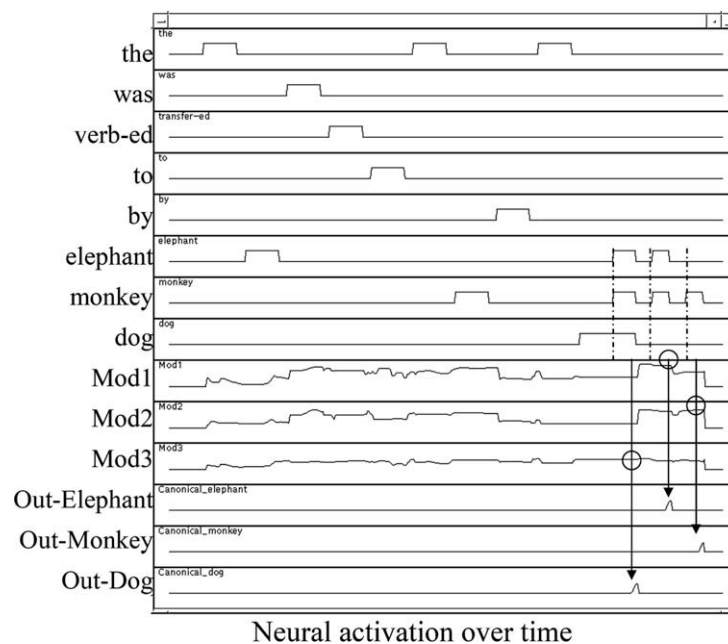


Fig. 2. Simulation example—comprehension of Dative Passive (non-canonical) sentence 'The elephant was transferred to the monkey by the dog.' In this simulation, the thematic role assignment takes place after the sentence has been presented, corresponding to a reanalysis and repair stage. Each row represents the activation of a given unit (neuron) in the model over time. Reading from left to right, the activation of these units reflects the presentation of the sentence to the model. Mod1–3 represent the activation of the three BA44/45M Modulation units that direct the contents of the first three BA45/47 elements, respectively, towards the output. Their activation is derived from inputs from BA47 based on learning. The 'Out' Rows represent the model's output responses corresponding to the three open-class units, always to be activated in the order agent, object, recipient for three-argument sentences. As soon as closed class words are presented, activity in BA47 (not shown) begins to drive the BA44/45M Mod1–3 neurons, based on the learning-modified BA47-Modulation synapses. When elephant, monkey and dog are all re-activated, this is the 'go signal' for the model to respond with the agent. At this point, the modulation neuron BA44/45M Mod3 that directs the contents of BA45/47(3) to Output is more active (indicated in circle on line Mod3) than Mod1 or Mod2 (813, 789 and 380, respectively), so the contents of BA45/47(3) (or 'dog') is directed towards the output to specify the agent, as seen in row N. Note the subsequent change in Mod1 and Mod2 activity, due to cortical input to basal ganglia (Eq. (6)) when the correctly chosen response matches with the corresponding element in the BA45/47 working memory. The remaining nouns (elephant and monkey) are then presented and the model should choose the object. The activity of the three Modulation units changes, and now BA44/45M Mod1 (1677) is more active than Mod2 (1369) or Mod3 (705). The contents of BA45/47(1) 'elephant' is thus directed to Output to specify the object. Finally, the remaining unassigned noun is presented and chosen to specify the recipient. Note that as required, BA44/45M Mod2 (1449) is more active than Mod1 (1172) and Mod3 (705). Thus, via the activation of appropriate BA44/45M neurons by BA47, the 3 nouns are selected in their canonical order, successfully analyzing the target sentence.

4.2. Syntactic comprehension results

Models instances were generated by setting the seed value of a random number generator used to create the synaptic strengths in the recurrent network. A set of 20 model subjects were isolated that were capable, after the training and testing described above, of processing all 6 two- and three-place verb sentences with no errors. Of these 20, 4 went on to successful completion of the 9 sentence types. Of these 4, 2 made the transition from 6 to 9 sentences with no additional training. That is, the capability to solve the 6 sentence configuration transferred directly to the nine sentence configuration. The remaining 2 subjects went on to master the 9 sentence version with no more than 40 exposures to each sentence type. Thus, 4 out of 20 subjects learned the complete Caplan task. This demonstrates that our hypothesis that the structure of closed-class elements forms unique identifiers for each sentence construction type holds. It also demonstrates that the recurrent cortical network can represent this structure.

It should be clearly stated that for this population that successfully mastered the nine sentence types, the learning produced a generalization capability that perfectly transfers to new sentences that follow one of the nine forms but use new nouns, or old nouns in a different order. That several distinct initial states (4/20) of the model architecture all converged on perfect performance of the nine sentence-type task indicates that the solution is a replicable property of the model architecture. This 'proof of concept' can be extended by the implementation of the recurrent network as a larger and higher dimensional complex system as demonstrated by Maas, Natschläger, and Markram (2002).

5. Artificial grammar learning

Syntactic comprehension and artificial grammar learning both share the underlying property that information contained within the sequence can be used to predict repetitive structural regularities in the rest of the sequence. In syntactic comprehension, the model should get the input sentence and then produce the open class elements in the canonical order. This repetition of the (possibly reordered) open class elements, constitutes the internal repetitive structure. The entire behavior (getting the sentence and then generating the canonically ordered open class elements) corresponds to a sequence with repetitive structure. In this case, the structure is encoded by the closed class elements. For artificial grammar learning, repetitive structure within the beginning of the sequence allows the system to predict the rest of the sequence, and thus to judge whether the whole sequence corresponds to a learned rule. In this context we can now apply the model to AGL task.

In studies of artificial grammar learning (AGL), subjects are exposed to strings of letters generated by a small finite-state grammar. These trained subjects are then able to classify new strings as grammatical or not at a level significantly above chance (Reber, 1967). In the changed letter set version, after initial training the subjects are asked to make the same kind of classification for strings generated (or not) from the same grammar, but using different letters (e.g., Gomez & Schvaneveldt, 1994). These subjects are capable of using grammatical knowledge acquired using one set of letters in order to classify strings generated (or not) from the same grammar but with different letters. Transfer tasks in AGL are a form of abstract sequence processing task (Dominey et al., 1998), with the major difference that classification ability is measure instead of reaction time. Thus, a model that performs the abstract sequencing tasks should also do well in AGL, and should fail in the same sense that humans fail. For example, in the transfer tasks, subjects that are trained on pairs of letters gen-

erated from the grammar perform significantly worse than those trained on longer strings of letters. (Gomez & Schvaneveldt, 1994), presumably because they do not have access to sufficient local context information in the strings.

To function in the changed letter set conditions, the system can only rely on information about relations between elements in strings. Thus, for example, a string ABABC has an internal repetitive structure that is preserved when the letters are changed as in DEFDEF (Dominey & Ramus, 2000; Dominey et al., 1998). The syntactic comprehension model in Fig. 1 is appropriate for this kind of task, as it has the necessary components of (1) a working memory to store the sequence as they are being processed, and (2) the capability to recognize repetition structure via the BA44/46R recognition system.

5.1. AGL simulation: transfer with changed letter sets in artificial grammar learning

This experiment simulated human performance in experiment 1 of Gomez and Schvaneveldt (1994). In that experiment, subjects were first exposed to strings (e.g., #WSWZ#, #WSSWZ#, ...) or letter pairs (#W, WS, SS, SW, ...) generated from a given grammar. They were then asked to classify another set of strings as grammatical or not. The purpose of the experiment was to determine if transfer of grammatical knowledge is facilitated if the initial training is with letter strings, vs. letter pairs, generated by the grammar. To determine if grammatical knowledge can be developed by exposure to pairs of letters (vs. strings of length 3 or more) generated by the grammar, humans were tested after training in these conditions. In the testing, subjects were required to classify new strings as legal or not based on the learned grammar. Two types of illegal strings were tested: strings with non-permissible pairs (NPP), and strings with permissible pairs in non-permissible locations (NPL). Pair training reduced the capacity to identify violations of non-permissible pairs, and eliminated the capacity to identify non-permissible location violations. We thus used two training conditions, and three testing conditions, applied to 10 'subjects,' created by using different seed values for the random number generator used to initialize the connection weights in the model. The 2 training conditions involved exposure to letter strings vs. pairs generated by the grammar G1 using letter set L1. The 3 testing conditions involved exposure to new strings using a different letter set from that used in training. These strings were either generated from the same training grammar (Gram), or violated the grammar by non-permissible pairs (NPP), or violated the grammar by having all legal pairs, but at least one of them in a non-permissible location (NPL). All training and testing items were taken exactly from Table 1 and Table 2 from Gomez and Schvaneveldt (1994). Letter strings or pairs in these tables were concatenated to form a sequence of elements that were then mapped onto elements in the 5×5 input array of the model, processed as open class elements in MTG.

Training consisted of exposure to the string (or pair) sequences, with six repetitions of the entire training set of 120 elements for a total of 720 responses. During testing, subsequent learning was prevented, and the trained model was exposed to the three testing conditions. For each testing condition, the sum of all response times was calculated and serves as a measure of grammaticality, with the value for Gram-L1 serving as a baseline against which the other values (NPP and NPL) can be compared.

5.2. AGL results

Fig. 3 summarizes the performance a model population ($n = 10$) on exposure to grammatical and non-grammatical strings generated (or not) by grammar G using new letter set L2. Values in the table are a measure of the mean and RTs for a sequence generated

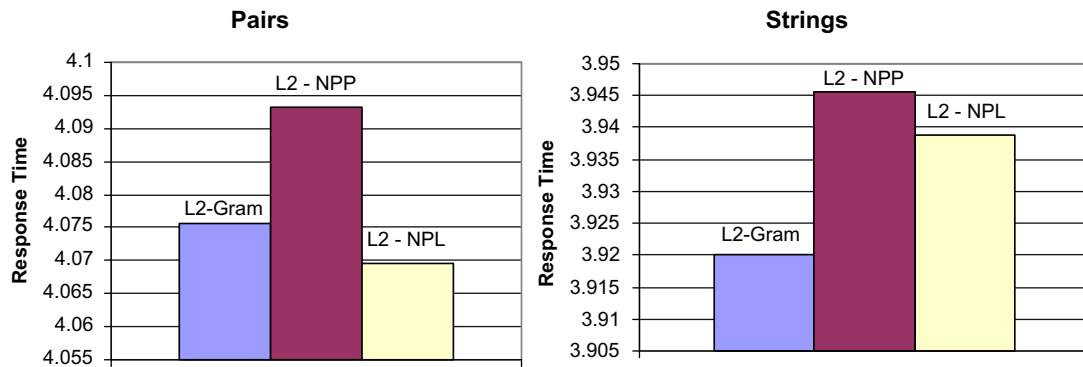


Fig. 3. Simulation performance for an artificial grammar task. Response times for grammatical (Gram), and non-grammatical (non-permissible pairs, NPP; and non-permissible location, NPL) strings, using the learned grammar, with a new letter set. A simulation time step is 0.005 s. Pairs training yields poor discrimination for NPP errors, and no discrimination for NPL errors, while the String training allows reliable discrimination for both.

by concatenating the test strings (or pairs). Increases in RT can be considered as an indication of a decrease in perceived grammaticality or predictability.

In the strings and pairs conditions the model is first retested with grammatical sequences (Gram), and then with the two types of non-grammatical strings, NPP and NPL. Fig. 3 illustrates that when the models are trained using letter strings, the discrimination capability for non-grammatical strings that have either violations of non-permissible pairs or non-permissible locations is sound as revealed by the increase in response times. When trained by letter pairs, however, the response for grammatical items is greater than that for string trained models, and the discrimination capability appears reduced. A repeated measures ANOVA revealed a significant effect for the training conditions (Strings superior to Pairs, $F(1,9) = 7.63$; $p = .02$), and a significant effect for testing conditions ($F(1,9) = 6.38$, $p = .008$), and most importantly, a significant interaction ($F(2,18) = 10.58$, $p < .001$). Post hoc analysis revealed that training with strings or pairs allowed a reliable discrimination of both type of errors. Importantly, string training allowed a significant capability to discriminate errors of the NPL type, whereas this discrimination capability was not significant when the model was trained with pairs, analogous to the behavior observed in human subjects (Gomez & Schvaneveldt, 1994).

The results indicate that the model again reliably captures several interesting aspects of human performance. Like humans, the model learns from a set of strings generated by a grammar using one letter set to identify at better than chance level new strings generated by the same grammar but with a new letter set.

6. General discussion

It seems intuitive that there should be a link between non-linguistic sequence processing and language processing. Marcus et al. (1999) examined infants' ability to recognize sequences with repetitive structures such as ABB vs. ABA and to transfer this knowledge to new sequences that had not been seen in training. This 'generative' ability to transfer to novel sequences seems a common characteristic with language processing. Dominey and Ramus (2000) simulated this performance with a model that had been developed indeed to explain the human ability to learn even more complex abstract sequences such as ABCBAC, DEFEDF, etc. both of which share the structure 123–213. The question still remained, what is the detailed relation between this type of sequence and language? We proposed a response, in which processing these abstract sequences involves a transformation from the initial triplet to the final triplet. Such a transformation could be analogous to a transformation from non-canonical order

to canonical order for sentence comprehension (Dominey et al., 2003). For example, consider: It was the ball¹ that John² gave to Bill³—John² gave the ball¹ to Bill³. Here the non-canonical relative form is transformed to the canonical form. Labeling the thematic roles gives the abstract representation 123–213. Thus, both abstract sequence processing and thematic role assignment can be characterized in terms of a mechanism that takes certain elements of an input string, and systematically transforms their order in the output.

We continued to explore this relation, and indeed, we found converging evidence that there was a neurophysiological common ground for sentence processing and this kind of abstract sequence processing (Hoen & Dominey, 2000; Hoen et al., 2003, 2006; Dominey et al., 2003). The underlying idea is that an encoding of sentence context (in terms of a small ensemble of possible cues as suggested by MacWhinney (1982) and Bates et al. (1982)) would then be used to retrieve the appropriate sentence-to-meaning mapping. In parallel a number of studies cited above provided detailed functional neurophysiology data on how these processes take place.

This provides the basis for the current research which represents the next level of progress and detail in establishing this link between linguistic and non-linguistic sequence learning. We believe that by exploring the question in this manner, we can make more significant progress in understanding the underlying neurophysiology. In this context, we have attempted to understand how the corticostriatal system plays a role in this processing.

The current research attempts to provide a 'single system' explanation for aspects of language acquisition and artificial grammar learning. Indeed, our simulation results demonstrate that a common model can (1) use patterns of closed class structure to learn the mappings of open class elements in sentences onto an ordered (canonical) representation of meaning in language, and (2) use patterns of repetitive structure to learn the mapping of elements in sequences onto a systematically reordered repetition of those elements further in the sequence. The common element, exploited in the model, is the ability to learn to recognize and manipulate regularities in structural mappings. In artificial grammars these regularities are found in the internal repetitive structure of the sequences. In language these regularities are found in the relation between word order and the canonical order in the meaning representation. In addition to learning the 9 sentence types from the Caplan test in English, we have examined the abilities of a related model which is based on the same underlying principals to acquire multiple languages. The underlying principal is that 'grammatical constructions as form to meaning mappings can be identified by the pattern of open and closed class elements

within a sentence of the language.' Dominey et al. (2006) thus demonstrated that a simplified version of the current model based on this principal could learn subsets of English, French and Japanese. In this context, second language acquisition should logically be accounted for as well. Perani and Abutalebi (2005) review data on L1 and L2 processing, and suggest a model in which a common neural system underlies acquisition of both. Differences in brain activity related to L1 vs. L2 processing are more related to computational load differences as a function of age of acquisition, proficiency, etc., with the underlying mechanisms remaining the same.

In the context of the degree of learning, the current model suggests a dual learning mechanism available for both sentence and non-linguistic sequence processing. A direct cortico-cortical mechanism may allow the use of syntactic structure information in BA47 to directly guide the modulation of semantic contents in BA45 into the unfolding representation of the sentence level meaning in BA44. In the case of conflict (e.g., detection that two elements are to be assigned to the same role), or excessive complexity, the system will resort to the long loop corticostriatal circuit, in order to retrieve a sentence level mapping rule that resolves the conflict. This rule will be more robust because it takes the entire sentence context into account. Ullman has proposed such a procedural system at the word-level, and we extend it here to encompass the sentence level (Ullman, 2004). We have implemented the corticostriatal mechanism here, and we have now specified an implementation strategy for the direct cortico-cortical mechanism.

In this context we can consider that as learning becomes established there will be the possibility of transfer of knowledge from the cortico-striatal loop to prefrontal cortex with extensive learning as suggested by Pasupathy and Miller (2005). These results support the hypothesis that rewarded associations are first identified by the basal ganglia, the output of which 'trains' slower learning mechanisms in the frontal cortex. In the context of language and artificial grammar processing, the cortico-striatal system would acquire the relations between context encoding and the appropriate transformations, which would require more extensive processing time. With over learning, these relations would be transferred to cortico-cortical encodings, which would allow for much more rapid online processing.

There remains much to be done in further understanding the functional neurophysiology of the LIFG and its relation with the basal ganglia. It appears clear that there is a processing gradient along Brodmann's areas 47, 45, 44, 6 and 9. Such gradients have been suggested by Hagoort (2005), and in a non-linguistic cognitive sequence processing context by Koehlin and Jubault (2006). While it is conceptually easier to partition specific roles for specific functions the possibility for overlap and combined functions remains a possibility. We tend to partition open and closed class processing such that grammatical structure building is guided only by closed class elements. However, in reality open class elements also contribute to structure building, and we can thus image connections in which semantic encoding in BA45 can also have an influence on the context encoding in BA47.

Finally, in the simulation of syntactic comprehension illustrated in Fig. 2, the thematic role assignment occurred after the entire sentence had been processed. This corresponds to a situation of reanalysis in which all information is available and the correct analysis can be performed. What about online assignment? We have shown in the artificial grammar learning task, that the model is capable of applying context information to the processing of elements as they arrive in the sequence on-line in real time (see also Dominey et al., 1998). This same capability will be effective in sentence processing, and indeed, should commit the same kind of early assignment errors for ambiguous sentence that people do. Recent neuroimaging data provide increasing detail on the time-

course of comprehension (Humphries, Binder, Medler, & Liebenthal, 2007). Our current research is addressing the mechanisms for online detection of errors and the neural mechanisms for changing the processing mode from on-line to post hoc reanalysis. In this context, Fig. 1 displays two pathways by which context information encoded in BA47 can influence the binding of semantic working memory contents in BA45 and BA45/47 into the appropriate thematic roles in BA46/6. A short cortico-cortical path by which BA47 influences BA44/45M directly, and a long loop by which BA47 activates striatum to retrieve the appropriate binding information at the sentence level via modifiable corticostriatal synapses and the basal ganglia loop. The localization of these distinct functional properties of different cortical regions in our model thus serve as the basis for future brain imagery experiments.

Appendix A. Specification of the corticostriatal model for syntactic comprehension and non-linguistic sequence learning

The model in Fig. 1 is implemented in Neural Simulation Language (NSL 2.1, Weitzenfeld, Arbib, & Alexander, 2002). Words are presented in the 25 element (5×5) Input array (not shown). Elements 0–12 are reserved for closed class words (one per element) and are directed to the closed class stream (STG). Elements 13–24 are reserved for open class words and are directed towards the open class stream (MTG). Each of the labeled structures corresponds to a 5×5 layer of leaky integrator neurons which represents the membrane potential of a population of neurons. Simulation time steps correspond to 5 ms real time. To generate the neural output firing rate, the leaky integrator value is passed through a sigmoid function. In the equations below the leaky integrator and sigmoid functions are represented together as the function $f(\cdot)$. More details of a related system can be found in Dominey et al. (1998, 2003).

Recurrent BA47 representation for the closed class stream: Eq. (1) describes how BA47 is influenced by closed class words from STG, recurrent inputs, and recognition of the transformed open class words from BA44/45R.

$$BA47 = f(STG * W^{STG-47} + BA47_D * W^{47-47} + \phi BA44/45R * W^{R-47}) \quad (1)$$

$$BA47_D = f(BA47) \quad (2)$$

The connections W^{STG-47} , W^{47-47} and W^{R-47} define the projections from units in STG, BA47D, and BA44/45R to BA47. Note that the latter has a modulation term ϕ which can vary between 1 and 0 depending on the task. 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 BA47 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 BA47 internal state information. Recurrent input to BA47 originates from the layer BA47D. BA47D (Eq. (2)) receives input from BA47, and its 25 leaky integrator neurons have a distribution of time constants from 100 to 2100 ms (20 to 420 simulation time steps), while BA47 units have time constants of 10 ms (2 simulation time steps). This distribution of time constants in BA47D yields a range of temporal sensitivity similar to that provided by using a distribution of temporal delays. The recurrent BA47 network thus maintains a representation of the context of closed class words that will be used to encode the structure of the sentence over time.

Open class stream: In order to represent and execute serial order transformations a system must (1) store the elements to be transformed in an accessible manner, and (2) provide a reliable method to access these elements in an order that may vary from the initial input order. The model of Fig. 1 realizes the first requirement with

a continuously updated working memory (BA45/47) of the open class elements. Each time an open class element enters the open class stream via MTG, it is held for immediate use in BA45 (Eq. (3.1)), and working memory in BA45/47 is updated, as described in Eqs. (3.2), (3.2) so that BA45/47 always contains the open class elements of the current sentence in the order they were presented. Each of the 5 BA45/47 elements is thus a 5×5 array. This working memory allows the model to perform reassignment of thematic roles after the sentence has been presented, e.g., in the case that initial assignments must be repaired.

$$\text{BA45} = \text{MTG} \quad (3.1)$$

$$\text{BA45}/47(1) = \text{BA45} \quad (3.2)$$

$$\text{BA45}/47(n) = \text{BA45}/47(n+1) \quad (3.3)$$

Learning: The objective of learning is that when open class elements arrive in BA45 (and/or in BA45/47), they can be modulated into the correct thematic role in BA44/6. This will occur as the sentence level coding of closed class structure in BA47 activates the caudate and retrieves the associated structure integration signal that activates the modulation neurons in BA44/45M. This results in binding of the open class element in BA45 to its correct thematic role in BA44/6. During learning, we consider that the meaning representation in BA44/6 is provided via visual perception of the scene that is being described in the sentence. This provides a form of supervised learning as the current open class element in BA45 can be compared with the contents of BA44/6 to determine its thematic role. The result of this comparison is stored in a 6-element vector called BA44/45R. Each BA44/45R element i , for $1 < i < 5$, is either zero if BA45 is different from BA44/6(i) or 1 if they are the same, as described in Eq. (4). The BA44/45R vector plays an important role, both in providing information to BA47 as described in Eq. (1), and also in the learning procedure itself.

$$\text{BA44}/45\text{R}(i) = \text{BA44}/6(i) * \text{BA45} \quad (4)$$

When a given Open class word such as 'Bill' is being processed in BA45, the recognition process BA44/45R detects a match with BA44/6(4) and projects this to the Caudate. This information traverses the corticostriatal system, and activates the corresponding BA44/45M modulation neurons that will modulate the contents of BA45 into BA44/6. Importantly, this assumes a preserved 'topography' such that activation of a given population in BA44/45R traverses the corticostriatal system to activate the corresponding element in BA44/45M. That is, recognition of a match will lead to generation of that same match in the future. At the same time, the BA47 context information is also projecting to caudate. Match detection triggers an internal dopamine reward-based learning (Eq. (5)), and, the BA47 activity becomes linked to activation of these same Caudate neurons.

BA44/45R encodes the match between the current open class element and its role in the meaning in BA44/6 during teaching. This match information can be used by BA44/45M to bind the current open class element onto its correct role in the meaning. This is how the comprehension process works when the meaning is not provided in advance. At the same time, BA47 encodes the current context. The goal of learning is that this context information becomes linked to the match information, so that it can be used to enforce the correct role binding. To exploit this information in the BA47 in order to bind the current open class element in BA45 to its appropriate role in BA44/6, the appropriate modulation neurons BA44/45M must be active. In order to achieve this, a new learning rule is developed. Once the learning has occurred, for a new sentence when the repetition is recognized between BA45 and an element in BA44/6(i) (BA44/45R, Eq. (4)), this information is transmitted to Caudate, selectively activating neurons that lead

to activation of the corresponding modulation neuron in BA44/45M via the basal ganglia loop.

The crucial learning occurs as follows in order to bind the BA47 activity to activation of this modulation unit. In Caudate, connections are strengthened between the active context (BA47) units and the Caudate units that are coactivated by the BA44/45R. The result is that this BA47 pattern becomes increasingly associated with the occurrence of the matched element such that, after learning, this association can be used to modulate or bind BA45 contents into the correct BA44/6 role via activation of BA44/45M units. This learning rule which modifies the corticostriatal synapses from BA47 is described in Eq. (5).

$$W_{ij}^{47-CD}(t+1) = W_{ij}^{47-CD}(t) + \text{BA47}_i * \text{BA44}/45\text{R}_j \quad (5)$$

The goal of this learning is to allow BA47 to modulate or bind the current BA45 element into BA44/6 to the correct thematic role. To permit this, a 5-element vector, BA44/45M, is introduced such that for $i = 1-5$, if BA44/45M(i) is non-zero, then the contents of BA45 is modulated or directed to BA44/6(i) (Eq. (6)). BA44/45M receives input from the corticostriatal system via Eq. (6.1).

$$\text{BA44}/45\text{M} = f(\text{Caudate}/\text{SNr}/\text{Thalamus}) \quad (6.1)$$

$$\text{Caudate}/\text{SNr}/\text{Thalamus} = f(\text{BA44}/45\text{R} + \text{BA47} * W^{47-CD}) \quad (6.2)$$

Based on the learning in Eq. (5), BA47 now directs this modulation of the BA45 contents into the BA44/6 meaning representation via BA47's influence on BA44/45M, as described in Eq. (6), with the resulting modulation or binding of the appropriate BA45 element into the output as described in Eq. (7). Note in Eq. (7) that the contents for the binding into BA44/6 can originate either from the current open class element in BA45, or from the working memory in BA45/47, which will be used in the case of late repair.

$$\text{BA44}/6(i) = \delta \text{BA45} * \text{BA44}/45\text{M}(i) + \varepsilon \text{BA45}/47 * \text{BA44}/45\text{M}(i) \quad (7)$$

After training on sequence 'The elephant was given to the monkey by the rabbit. - rabbit, elephant, monkey,' when the model is exposed to a new isomorphic sequence 'The giraffe was given to the lion by the tiger' and is processing the open class word 'lion,' the closed class sequence 'the _ was _en to the _' will evoke a pattern of activation of BA47 units that will drive the caudate and the corresponding modulation BA44/45M(3) element, thus directing 'lion' in BA45 to its appropriate thematic role of Recipient in BA44/6. The changing BA47 activity will allow the corresponding assignments to be made for the agent and object roles. Errors will invoke the learning rule with a negative learning rate, thus reducing the probability of a repetition of the same error.

Of course, initial assignments for the first noun in a sentence will often be wrong, as they are underspecified before the entire sentence is seen. This is why we previously (Dominey et al. 2003, 2006) waited until the end of the sentence before assigning thematic roles, based on the Caplan task. The solution will be thus to employ both methods: A first pass assignment is made online—then as information accumulates, reassignments can be made. This implies (1) maintaining the open class elements in BA45/47 working memory, and (2) maintaining a sentence level mapping of each of these elements onto the BA44/6 representation, such that this mapping can potentially change during the course of the sentence processing. This is reflected in Eq. 7, where BA44/45M influences both BA45 for the online case, and the working memory of BA45/47 for the reanalysis case. Our current research addresses this.

A.1. Artificial grammar learning

We perform these experiments in a sequence learning context in which reaction times of the model are used as a performance measure. Because of the leaky integrators with threshold firing rates, the degree of input activation will influence the rate at which the network responds, and learning thus results in reduced reaction times for sequences that follow the learned structure (see Dominey et al., 1998). In this context we simply introduce an output component to the network that is influenced both by the input, and by the learning-related modulation of working memory contents (Eq. (8)). The idea is that the direct input to output connect will provide a baseline reaction time, and that the learning related modulation will provide for reduced reaction times.

The working memory in BA45/47 invariantly remember the 7 previous sequence elements. The recognition function BA44/45R compares the current response with the contents of these BA45/47 elements (Eq. (4)*). BA47 receives input from the recognition function and thus encodes the abstract repetitive structure of the sequence (Eq. (1)). The modulatory (BA44/45M) neurons feed the BA45/47 representations into the output (Eq. (8)). This modulation is under the control of BA47 (Eq. (6.1)*) which has now been modified to include both the direct cortico-cortical path from BA47, and the long basal ganglia path via caudate, SNr and thalamus.

Each time a recognition occurs the current context activity in BA47 (which now encodes the repetitive structure of the sequence) becomes associated with the neuron BA44/45M(i) that modulates the contents of the recognized BA45/47 element BA45/47_i to Out (Eq. (8)). The next time we arrive at this same point in a sequence with the same structure, the same state will be encoded in BA47 so it will modulate the contents of BA45/47_i into the output structure Out (via the recently updated connections in W_{ij}^{47-M} , from Eq. (9)) thus allowing for prediction and reduced RT for that response. After training on ABABC, the model will recognize (and respond with reduced RTs) all sequences with an abstract structure of the form 1–2–1–2–3 (or 'unique, unique, match $n - 2$, match $n - 2$, unique' in terms of recognitions detected by BA44/45R). Thus, in addition to learning a sequence of surface elements, the model learns a sequence of abstract structure 'slots' that can have arbitrary 'fillers,' thus allowing learning of the internal repetitive structure of artificial grammars.

$$\text{Out} = f(\text{BA45} + \text{BA44/45M} * \text{BA45/47}) \quad (8)$$

$$\text{BA44/45R}(i) = \text{BA45/47}(i) * \text{BA45} \quad (4^*)$$

$$\text{BA44/45M} = \alpha \text{Caudate/SNr/Thalamus} + \gamma \text{BA47} * W^{47-M} \quad (6.1^*)$$

$$W_{ij}^{47-M}(t+1) = W_{ij}^{47-M}(t) + \text{BA47}_i * \text{BA44/45R}_j \quad (9)$$

References

- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9, 357–381.
- Alexander, M. P., Naeser, M. A., & Palumbo, C. L. (1987). Correlations of subcortical CT lesion sites and aphasia profiles. *Brain*, 110, 961–991.
- Bahlmann, J., Gunter, T. C., & Friederici, A. D. (2006). Hierarchical and linear sequence processing: An electrophysiological exploration of two different grammar types. *Journal of Cognitive Neuroscience*, 18(11), 1829–1842.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11, 245–299.
- Behrens, T. E., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A., Boulby, P. A., et al. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6(7), 750–757.
- Calabresi, P., Picconi, B., Tozzi, A., & Di Filippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends in Neurosciences*, 30(5), 211–219.

- Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, 21, 117–175.
- Caplan, D., Stanczak, L., & Waters, G. (2008). Syntactic and thematic constraint effects on blood oxygenation level dependent signal correlates of comprehension of relative clauses. *Journal of Cognitive Neuroscience*, 20(4), 643–656.
- De Witte, L., Wilssens, I., Engelborghs, S., De Deyn, P. P., & Mariën, P. (2006). Impairment of syntax and lexical semantics in a patient with bilateral paramedian thalamic infarction. *Brain and Language*, 96(1), 69–77.
- Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., et al. (2008). Functional connectivity of human striatum: A resting state fMRI study. *Cerebral Cortex* [pub ahead of print].
- Dominey, P. F. (2005a). From sensorimotor sequence to grammatical construction: Evidence from simulation and neurophysiology. *Adaptive Behavior*, 13(4), 347–362.
- Dominey, P. F. (2005b). Emergence of grammatical constructions: Evidence from simulation and grounded agent experiments. *Connection Science*, 17(3–4), 289–306.
- Dominey, P. F., & Arbib, M. A. (1992). A cortico-subcortical model for generation of spatially accurate sequential saccades. *Cerebral Cortex*, 2(2), 153–175.
- 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., Hoen, M., Blanc, J.-M., & Lelekov-Boissard, T. (2003). Neurological basis of language and sequential cognition: Evidence from simulation, aphasia and ERP studies. *Brain and Language*, 86, 207–225.
- Dominey, P. F., Hoen, M., & Inui, T. (2006). A neurolinguistic model of grammatical construction processing. *Journal of Cognitive Neuroscience*, 18(12), 2088–2107.
- Dominey, P. F., 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, 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. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71–99.
- Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, 24, 79–91.
- Fletcher, P., Büchel, C., Josephs, O., Friston, K., & Dolan, R. (1999). Learning-related neuronal responses in prefrontal cortex studies with functional neuroimaging. *Cerebral Cortex*, 9(2), 168–178.
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage*, 32, 956–967.
- 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. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84.
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Science of the United States of America*, 103, 2458–2463.
- Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *NeuroImage*, 20, S8–S17.
- Friederici, A. D., Rüschemeyer, S. A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb Cortex*, 13(2), 170–177.
- Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In B. Mountcastle Vernon (Ed.). *Handbook of physiology* (vol. 5, pp. 373–417). Bethesda, MD: American Psychological Society.
- Gomez, R. L., & Schvaneveldt, R. W. (1994). What is learned from artificial grammars? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 396–410.
- Grossman, M. (1999). Sentence processing in Parkinson's disease. *Brain and Cognition*, 40, 387–413.
- Hagoort, P. (2005). On Broca, brain and binding: A new framework. *Trends in Cognitive Science*, 9, 416–423.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2), 301–307.
- Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension in Parkinson's disease. *Brain and Language*, 97, 243–257.
- 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.
- Hoen, M., Golembiowski, M., Guyot, E., Deprez, V., Caplan, D., & Dominey, P. F. (2003). Training with cognitive sequences improves syntactic comprehension in agrammatic aphasics. *NeuroReport*, 14(3), 495–499.
- Hoen, M., Pachot-Clouard, M., Segebarth, C., & Dominey, P. F. (2006). When Broca experiences the Janus syndrome: An ER-fMRI study comparing sentence comprehension and cognitive sequence processing. *Cortex*, 42(4), 605–623.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *NeuroImage*, 36(3), 924–932.

- Inui, T., Ogawa, K., & Ohba, M. (2007). Role of left inferior frontal gyrus in the processing of particles in Japanese. *NeuroReport*, 18(5), 431–434.
- Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, 50, 963–974.
- Leh, S. E., Pfito, A., Chakravarty, M. M., & Strafella, A. P. (2007). Fronto-striatal connections in the human brain: A probabilistic diffusion tractography study. *Neuroscience Letters*(419), 113–118.
- Léhéricy, S., Ducros, M., Krainik, A., Francois, C., Van de Moortele, P.-F., Ugurbil, K., et al. (2004). 3-D diffusion tensor axonal tracking shows distinct SMA and pre-SMA projections to the human striatum. *Cerebral Cortex*, 14, 1302–1309.
- Maas, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14, 2531–2560.
- MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Language development, syntax and semantics* (Vol. 1, pp. 73–136). Hillsdale, NJ: Erlbaum.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77–80.
- Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., & Fazio, F. (2001). Syntax and the brain: Disentangling grammar by selective anomalies. *NeuroImage*, 13, 110–118.
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., & Buchel, C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, 6, 774–781.
- 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 & Performance*, 24(3), 1–11.
- Ogawa, K., Ohba, M., & Inui, T. (2007). Neural basis of syntactic processing of simple sentences in Japanese. *NeuroReport*, 18(14), 1437–1441.
- 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.
- Pasupathy, A., & Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*, 433, 873–876.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202–206.
- Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, 28, 383–407.
- Pickett, E. R., Kuniholm, E., Protopapas, A., Friedman, J., & Lieberman, P. (1998). Selective speech motor, syntax and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: A case study. *Neuropsychologia*, 36(2), 173–188.
- Postuma, R. B., & Dagher, A. (2006). Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cereb Cortex*, 16(10), 1508–1521.
- Pulvermüller, F. (1995). Agrammatism: Behavioral description and neurobiological explanation. *Journal of Cognitive Neuroscience*, 7(2), 165–181.
- Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *The European Journal of Neuroscience*, 21(3), 793–807.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Sakai, K. L. (2005). Language acquisition and brain development. *Science*, 403, 815–819.
- Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain Lang*, 52, 452–473.
- Teichmann, M., Dupoux, E., Kouider, S., Brugieres, P., Boisse, M. F., Baudic, S., et al. (2005). The role of the striatum in rule application: The model of Huntington's disease at early stage. *Brain*, 128(Pt 5), 1155–1167.
- Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., & Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *NeuroImage*, 17, 700–709.
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17(2), 273–281.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92, 231–270.
- Ullman, M. T. (2006). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex*, 42(4), 480–485.
- Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., et al. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9, 266–276.
- Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*(6), 131–138.
- Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*(37), 159–170.
- Weitzenfeld, A., Arbib, M., & Alexander, A. (2002). *NSL—Neural simulation language: A system for brain modeling*. MIT Press.
- Wiesendanger, E., Clarke, S., Kraftsik, R., & Tardif, E. (2004). Topography of cortico-striatal connections in man: Anatomical evidence for parallel organization. *The European Journal of Neuroscience*(20), 1915–1922.