Implicit Artificial Syntax Processing: Genes, Preference, and Bounded Recursion

Vasiliki Folia, Christian Forkstam, Martin Ingvar,
Peter Hagoort & Karl Magnus Petersson

The first objective of this study was to compare the brain network engaged by preference classification and the standard grammaticality classification after implicit artificial syntax acquisition by re-analyzing previously reported event-related fMRI data. The results show that preference and grammaticality classification engage virtually identical brain networks, including Broca's region, consistent with previous behavioral findings. Moreover, the results showed that the effects related to artificial syntax in Broca's region were essentially the same when masked with variability related to natural syntax processing in the same participants. The second objective was to explore CNTNAP2-related effects in implicit artificial syntax learning by analyzing behavioral and event-related fMRI data from a subsample. The CNTNAP2 gene has been linked to specific language impairment and is controlled by the FOXP2 transcription factor. CNTNAP2 is expressed in language related brain networks in the developing human brain and the FOXP2–CNTNAP2 pathway provides a mechanistic link between clinically distinct syndromes involving disrupted language. Finally, we discuss the implication of taking natural language to be a neurobiological system in terms of bounded recursion and suggest that the left inferior frontal region is a generic on-line sequence processor that unifies information from various sources in an incremental and recursive manner.

Keywords: artificial grammar learning; artificial language; Broca's region; CNTNAP2; fMRI; FOXP2; genes, grammaticality classification; natural language; preference classification; syntax

1. Introduction

Human languages are characterized by universal "design features" (Hockett 1963, 1987): discreteness, arbitrariness, productivity, and the duality of patterning

This work was supported by the Max Planck Institute for Psycholinguistics, the Donders Institute for Brain, Cognition and Behaviour, the Fundação para a Ciência e Tecnologia (PTDC/PSI-PCO/110734/2009; IBB/CBME, LA, FEDER/POCI 2010), the Stockholm Brain Institute, Vetenskapsrådet, the Swedish Dyslexia Foundation, the Hedlunds Stiftelse, and the Stockholm County Council (ALF, FoUU).

ISSN 1450–3417

(i.e. elements at one level are combined to construct elements at another). Somehow these properties arise from the way the human brain processes, develops, and learns, in interaction with its environment. The human capacity for language and communication is subserved by a network of brain regions that collectively instantiate the phonological, syntactic, semantic, and pragmatic operations necessary for adequate language production and comprehension. During normal language processing, phonology, syntax, and semantics operate in close temporal and spatial contiguity in the human brain. Therefore the artificial grammar learning (AGL) paradigm has been used to create a relatively uncontaminated window onto the neurobiology of syntax. Artificial syntax learning paradigms thus makes it possible to investigate structured sequence processing relatively independent of, for example, semantics and phonology (Petersson *et al*. 2004, 2010). In addition, artificial syntax learning has been used for cross-species comparisons in an attempt to establish the uniquely human component of the language faculty (Hauser *et al*. 2002, Fitch & Hauser 2004, O'Donnell *et al*. 2005, Gentner *et al*. 2006, Saffran *et al*. 2008).

Artificial syntax learning paradigms have been widely employed to study different aspects of natural language acquisition (Gómez & Gerken 2000, Folia *et al*. 2010), though it was originally implemented to investigate the underlying implicit sequence learning mechanism, which is presumably shared with natural language learning (Reber 1967) as well as other situations in which new skills are acquired (e.g. Misyak *et al*. 2009, 2010a, 2010b). The neurobiology of implicit sequence learning as assessed by artificial syntax acquisition have been investigated by means of functional neuroimaging (e.g. Petersson *et al*. 2004, 2010, Forkstam *et al*. 2006), brain stimulation (Uddén *et al*. 2008, 2011, de Vries *et al*. 2010), and agrammatic aphasics (Christiansen *et al*. 2010), and generally involve fronto-striatal circuits (Packard & Knowlton 2002, Ullman 2004; note that implicit learning is sometimes referred to as procedural learning, and vice versa), which are also involved in the acquisition of natural syntax (Ullman 2004). More specifically, recent functional neuroimaging (e.g. Petersson *et al*. 2004, 2010, Forkstam *et al*. 2006) and brain stimulation research (Uddén *et al*. 2008, 2011, de Vries *et al*. 2010), have identified some of the brain regions involved, including repeatedly showing that Broca's region, a brain region involved in natural syntax processing, is also involved in artificial syntax processing. Indeed, the breakdown of syntax processing in agrammatic aphasia is associated with impairments in artificial syntax learning (Christiansen *et al*. 2010). Moreover, Conway & Pisoni (2008) found that individual variability in implicit sequence learning correlated with language processing. Supportive evidence also comes from a recent study by Misyak *et al*. (2010a), who found that individual differences in learning non-adjacent dependencies, assessed by non-linguistic implicit sequence learning, correlate with the processing of natural language sentences containing complex non-adjacent dependencies. This supports the hypothesis that artificial grammar learning paradigm taps into implicit structured sequence learning and artificial syntax processing, and thus provides a useful way to investigate aspects of natural language processing. Thus, there is a growing body of evidence that language acquisition and language processing, both a natural and artificial setting, is mediated by implicit sequence learning and structured sequence proces-

sing mechanisms, respectively.

The implicit artificial syntax learning paradigm allows for a systematic investigation of aspects of structural acquisition from grammatical examples without providing explicit feedback, teaching instruction, or engaging the subjects in explicit problem solving (Forkstam *et al*. 2006, 2008, Folia *et al.* 2008). These acquisition conditions resemble, in certain important respects, those found in natural-language development with respect to syntax acquisition (Chomsky & Miller 1963: 275–276). Generally, artificial grammar learning paradigms consist of acquisition and test phases. In the acquisition phase, participants are exposed to an acquisition sample generated from a formal grammar. In the standard version, subjects are informed that the sequences were generated according to a complex set of rules after acquisition, and are asked to classify novel sequences as grammatical or not, based on their immediate intuitive impression (i.e. guessing based on gut-feeling). A well-replicated and robust finding in this paradigm is that subjects perform well above chance after several days of implicit acquisition; they do so on regular (e.g. Stadler & Frensch 1998, Folia *et al.* 2008, Forkstam *et al.* 2008) as well as non-regular grammars, including those that generate context-free and context-sensitive non-adjacent dependencies (Uddén *et al.* 2009).

In this study, we investigate an implicit preference AGL paradigm with several days of acquisition. During the implicit acquisition period, participants were exposed to grammatical sequences only in a cover task based on the structural mere-exposure effect (Zajonc 1968, Zizak & Reber 2004, Folia *et al.* 2008, Forkstam *et al.* 2008). The structural mere-exposure effect refers to the finding that repeated exposure to a stimulus created by a certain rule system, induces an increased preference for novel stimuli conforming to the same underlying system (Zizak & Reber 2004). To this end, we exposed the participants to a simple right-linear unification grammar — a grammar that generates right-linear phrase structures (Vosse & Kempen 2000, Hagoort 2005, Petersson *et al.* 2010). During the acquisition period, spanning five days, subjects were exposed to syntactically well-formed consonant sequences and no performance feedback was provided. On the last day a preference classification test was administered in which new sequences were presented. Previously, the implicit preference AGL paradigm has been characterized exclusively in behavioural terms (e.g. Manza & Bornstein 1995, Zizak & Reber 2004, Folia *et al.* 2008, Forkstam *et al.* 2008). Here we first review the outcome of implicit artificial syntax acquisition from an event-related fMRI study (Folia *et al*. 2011). Then we compare the brain network engaged by preference classification and the standard grammaticality classification after implicit artificial syntax acquisition from a previously reported event-related fMRI results on the standard grammaticality classification paradigm in the same subjects (Petersson *et al.* 2010). In addition, we investigate the common overlap between artificial and natural syntax processing by masking the non-grammatical (NG) vs. grammatical (G) effect observed in preference classification with the natural-syntax-related variability in the same subjects (Folia *et al*. 2009). Consistent with the hypothesis of implicit utilization of acquired structural knowledge as well as previous behavioral results (Forkstam *et al.* 2008), which showed that subjects perform qualitatively identical on preference and grammaticality classification, we found that the brain network subserving preference

classification during artificial syntax processing engaged Broca's region cantered on Brodmann's areas (BA) 44 and 45, and did not differ from those observed during grammaticality classification. This strengthens the notion that preference and grammaticality classification in the implicit artificial syntax learning are essentially equivalent (Forkstam *et al.* 2008). Finally, based on these event-related fMRI data (Petersson *et al*. 2010, Folia *et al*. 2011), we took advantage of the fact that a subsample of our participants was part of the Brain Imaging Genetics (BIG) project at the Donders Centre for Cognitive Neuroimaging and the Department of Human Genetics of the Radboud University Nij-megen. This allowed us to explore the potential role of the CNTNAP2 gene in artificial syntax acquisition/ processing at the behavioral as well as the brain level.

Relatively recently, language research has started to investigate the role of genes in language (Enard *et al.* 2002, Vargha-Khadem *et al*. 2005, Bishop 2009, Konopka *et al.* 2009). For example, mutations in the FOXP2 gene result in a complex symptomatology, called developmental verbal dyspraxia, which includes difficulties with learning and producing sequences of oral movements relevant for speech, as well as impairments in morphosyntactic aspects of language processing (Lai *et al*. 2001, Watkins *et al*. 2002, MacDermot *et al*. 2005). FOXP2 is a gene that codes for the transcription factor (a protein) foxp2 which regulates gene expression during development. This means that foxp2 controls the production of other proteins coded for by other genes. Transcription factors and their genes make up complex gene regulatory networks, which control many complex biological processes, including ontogenetic development (Davidson *et al*. 2002, Davidson 2006, Alberts *et al.* 2007). Moreover, functional neuroimaging studies of the KE family (with a protein-truncating FOXP2 mutation; Lai *et al.* 2001), have demonstrated structural and functional abnormalities in brain regions related to language (Vargha-Khadem *et al.* 2005). The CNTNAP2 gene has been linked to specific language impairment (SLI) and the FOXP2–CNTNAP2 pathway provides a mechanistic link between clinically distinct syndromes involving disrupted language (Vernes *et al.* 2008). The CNTNAP2 gene is controlled (down-regulated) by the foxp2 transcription factor (Vernes *et al.* 2008). CNTNAP2 codes for a neural trans-membrane protein, which belongs to neurexin superfamily (Poliak *et al.* 1999) and it has been shown that, in the developing human brain, the expression of CNTNAP2 is relatively increased in fronto-temporal-subcortical brain networks (Alarcón *et al.* 2008). In particular, the CNTNAP2 expression is enriched in frontal brain regions in humans, but not in mice or rats (Abrahams *et al.* 2007). A recent study investigated the effects of a common single nucleotide polymorphism (SNP) RS7794745 in the CNTNAP2 gene on the brain response during language comprehension (Snijders *et al*. 2011). This study found both structural and functional brain differences in language comprehension related to the same SNP sub-grouping used in this study.

Finally, we note that an artificial grammar represents a formal specification of the mechanism that generates, for example, specific structural or sequence regularities (e.g., various types of local or non-adjacent dependencies). From this point of view, an artificial syntax is a formal language (Davis *et al*. 1994) and artificial syntax learning is an experimental model to investigate various (any) generative mechanism independent of other aspects of a language (cf. the

introduction of Petersson *et al.* 2004). As noted above, artificial syntax learning can be used as an experimental tool to investigate the processing properties of Broca's region, a central node in the brain network for natural syntax processing. In this context, we take the view that natural and artificial syntax processing share a common abstraction — structured sequence processing. Clearly, any particular artificial grammar cannot instantiate all phenomena found in natural syntax. Rather, in experimental work it is necessary to focus on some particular aspect of syntax, which is also the case for experimental work on natural language syntax. Artificial syntax learning thus provides a window onto the neurobiology of syntax, in the sense that artificial syntax learning allows us to investigate the computational properties of Broca's region. In the Discussion section, we return to some issues related to the present the Chomsky hierarchy and recursive processing from the point of view that natural language is a neurobiological system.

2. Materials and Methods

2.1. Participants

Here we briefly describe the relevant background of the material and methods used by Vasiliki Folia and colleagues (Folia *et al*. 2008, 2011, Petersson *et al*. 2010) as they apply to this study. Thirty-two healthy right-handed Dutch university students were recruited in the study (16 females, mean age ± SD = 22 ± 3 years; mean years of education ± SD = 16 ± 2). None of the subjects used any medication, had a history of drug abuse, head trauma, neurological or psychiatric illness, or a family history of neurological or psychiatric illness. All subjects had normal or corrected-to-normal vision. Written informed consent was obtained from all participants according to the Declaration of Helsinki as well as from the local medical ethics committee. Of the thirty-two participants, twelve were already included in the BIG database at the Donders Centre for Cognitive Neuroimaging and the Department of Human Genetics of the Radboud University Nijmegen (5 females, mean age ± SD = 22 ± 2 years; mean years of education ± SD = 16 ± 2) and typed for the single nucleotide poly-morphism (SNP) RS7794745 (with a breakdown on AA:AT:TT of 4:6:2). Because of the few TT-carriers, we pooled all T-carriers into one group of TT- and AT-carriers and analyzed the data in the T (N = 8) and nonT (N = 4) groups.

2.2. Stimulus Material

We used a simple right-linear unification grammar (Petersson *et al.* 2010) with the following vocabulary of terminal symbols (M, S, V, R, X) and lexicon of primitive trees (*treelets*) {[s_1, [M, s_2]], [s_2, [S, s_2]], [s_2, [V, s_4]], [s_3, [X, s_2]], [s_3, [X, s_5]], [s_4, [R, s_3]], [s_4, [S, s_6]], [s_4, #], [s_5, [R, s_5]], [s_5, [M, s_6]], [s_5, #], [s_6, #]}. For a given lexical item (e.g., [s_j, [T, s_k]]), s_j, s_k can be interpreted as syntactic control features and T as a surface feature. Within the unification framework (Vosse & Kempen 2000, Hagoort

2005, Petersson *et al.* 2010), an incoming sequence of surface symbols (e.g., MSV) initiates the retrieval of lexical items from the mental lexicon. As a result, they enter a unification space for on-line processing: [s_1, [M, s_2]], [s_2, [S, s_2]], [s_2, [V, s_4]]…, where two lexical items (e.g., [s_i, [R, s_j]], [s_k, [Q, s_l]]) unify (i.e. combine or merge) through a *unification operation* U if and only if $s_j = s_k$, or $s_l = s_i$. This process is *incremental and recursive*. For example, if the structure U([s_1, [M, s_2]], [s_2, [S, s_2]]) = [s_1, [M, [s_2, [S, s_2]]]] is already present in the unification space when the lexical item [s_2, [V, s_4]] is retrieved, a larger combinatorial structure can be formed by the unification operation U([s_1, [M, [s_2, [S, s_2]]]], [s_2, [V, s_4]]) = [s_1, [M, [s_2, [S, [s_2, [V, s_4]]]]]], and so on. The Unification operator works in the same way in all unification grammars. However, the structures generated by the Unification operator depend on the structure of the lexical items in any given grammar. In the present case, our grammar yields right-linear structures.

Folia *et al.* (2011) used a 2 x 2 x 2 factorial design including the factors instruction type (preference/grammaticality instruction), grammaticality status (grammatically correct/incorrect), and local subsequence familiarity (high/low ACS). The local subsequence familiarity (cf. Knowlton & Squire 1996, Meulemans & van der Linden 1997, Forkstam *et al.* 2006 for technical descriptions) is an associative measure of the superficial resemblance between classification sequences and the sequences in the acquisition set. The classification sequences with high ACS contain subsequences (bigrams and trigrams) that appear frequently in the acquisition set, while sequences with low ACS contain subsequences with a low frequency in the acquisition set. In total, 569 G sequences from the grammar, with a sequence length ranging from 5 to 12, were generated. For each item the frequency distribution of 2 and 3 letter chunks for both terminal and complete sequence positions was calculated. In this way, the associative chunk strength (ACS) was calculated for each item (cf. Knowlton & Squire 1996, Meulemans & van der Linden 1997, Forkstam *et al.* 2006). Next, for the acquisition set, 100 sequences representative, in terms of letter chunks, for the complete sequence set were randomly selected in an iterative way. In the next step, the NG sequences were created, derived from non-selected G sequences, by switching letters in two non-terminal positions. The NG sequences matched the G sequences in terms of both terminal and complete-sequence ACS (Forkstam *et al.* 2006, 2008). Finally, in an iterative procedure, we randomly selected two sets of 56 sequences each from the remaining G sequences, to serve as classification sets. The classification sets thus consisted of 25% grammatical/high ACS (HG); 25% grammatical/low ACS (LG); 25% non-grammatical/high ACS (HNG); and 25% non-grammatical/low ACS (LNG) sequences. See Appendix A below for example stimuli.

2.3. Experimental Procedures

During the acquisition sessions, subjects were presented with the 100 acquisition sequences (presentation order randomized for each acquisition session) and the task was an immediate short-term memory task serving as a cover task. Each sequence was centrally presented letter-by-letter on a computer screen (3–7 s corresponding to 5–12 terminal symbols; 300 ms presentation, 300 ms inter-symbol-interval) using the Presentation software (http://nbs.neuro-bs.com).

When the last letter in a sequence disappeared, subjects were instructed to reconstruct the sequence from memory and type it on a keyboard. No performance feedback was given, and only grammatical sequences were presented. The acquisition phase lasted approximately 20–40 minutes and took place over five consecutive days.

After the acquisition session on the last (5th) day of the experiment, subjects participated in a preference and then a grammaticality classification session. During preference classification, subjects were presented with new sequences, which they have not seen before. They were instructed to classify the new sequences according to their immediate intuitive preference (i.e. guessing whether they liked the sequence, or not, based on gut-feeling; *preference instruction*). Subsequently, they were informed about the existence of a generating set of rules and the subjects were asked to classify new sequences as grammatical or not based on their gut-feeling (*grammaticality instruction*). fMRI data were acquired during both preference and grammaticality classification (Petersson *et al*. 2010, Folia *et al*. 2011).

The classification sequences were presented via an LCD-projector on semi-transparent screen that the subject comfortably viewed through a mirror mounted on the head-coil. The classification sessions were split in two parts, in order to balance response finger within subjects (subjects indicated their decision by pushing the corresponding response key with their left/right index finger). Each part lasted approximately 20 minutes. After a 1 s pre-stimulus period, the sequences were presented sequentially, followed by a 3 s response window. A low-level baseline condition was also included; a sensorimotor decision task in which sequences of letters P or L (matched for sequence length to the classification set) were presented in the same fashion as the classification sequences and subjects responded by pressing the right or left index finger, respectively. The different sequence types were presented in random order.

3. Data Acquisition and Statistical Analysis

Behavioral data were analyzed with repeated measures ANOVAs (SPSS 15.0) with non-sphericity correction. A significance level of $P < .05$ was used throughout. Data analysis was carried out for the whole group and the sub-sample for which CNTNAP2 (SNP RS7794745) data were available (T-group: AT/TA/TT allele; nonT-group: AA allele).

3.1. MR Data Acquisition

Whole head T2*-weighted functional echo planar blood oxygenation level dependent (EPI-BOLD) fMRI data were acquired with a Siemens Avanto 1.5T scanner using an ascending slice acquisition sequence (volume TR = 2.6s, TE = 40 ms, 90 degree flip-angle, 33 axial slices, slice-matrix size = 64x64, slice thickness = 3 mm, slice gap = .5 mm, FOV = 224 mm, isotropic voxel size = 3.5x3.5x3.5 mm^3) in a randomized event related fashion. For the structural MR image volume, a high-resolution T1-weighted magnetization-prepared rapid gradient-echo pulse

sequence was used (MP-RAGE; volume TR = 2250 ms, TE = 3.93 ms, 15 degree flip-angle, 176 axial slices, slice-matrix size = 256x256, slice thickness = 1 mm, field of view = 256 mm, isotropic voxel-size = 1.0x1.0x1.0 mm^3).

3.2. MR Image Pre-Processing and Statistical analysis

We used the SPM5 software for image pre-processing and statistical analysis. The EPI-BOLD volumes were re-aligned to correct for individual subject movement and were corrected for differences in slice acquisition time. The subject-mean EPI-BOLD images were subsequently spatially normalized to the functional EPI template provided by SPM5. The normalization transformations were generated from the subject-mean EPI-BOLD volumes and applied to the corresponding functional volumes. The functional EPI-BOLD volumes were transformed into the MNI space, an approximate Talairach space (Talairach & Tournoux 1988), defined by the SPM5 template, and spatially filtered with an isotropic 3D spatial Gaussian kernel (FWHM = 10 mm). The fMRI data were analyzed statistically, using the general linear model framework and statistical parametric mapping in a two-step random-effects summary-statistics procedure (Friston *et al*. 2007). We included the realignment parameters for movement artifact correction and a temporal high-pass filter (cycle cut-off at 128 s), to account for various low-frequency effects.

At the first-level, single-subject analyses were conducted. The linear model included explanatory regressors modeling the sequence presentation period from the position of the anomaly in the HNG and LNG conditions and their correct counterparts in the HG and LG conditions. This was done separately for correct and incorrect responses. The initial part of the sequences was modeled separately, as was the baseline and the inter-sequence-interval. The explanatory variables were temporally convolved with the canonical hemodynamic response function provided by SPM5. At the second-level, we generated single-subject contrast images for the correctly classified HG, LG, HNG, and LNG sequences, relative to the sensorimotor decision baseline. These were analyzed in a random-effects repeated-measure ANOVA with non-sphericity correction for repeated measures and unequal variance between conditions. Statistical inference was based on the cluster-size test-statistic from the relevant second-level SPM[T] maps thresholded at P = .005 (uncorrected). Only clusters significant at $P_{FWE} < .05$ family-wise error (FWE) corrected for multiple non-independent comparisons, based on smooth random field theory (Adler 1981, Worsley *et al.* 1996, Adler & Taylor 2007, Friston *et al.* 2007) are described. In addition, we list the coordinates of local maxima and their corresponding P-values corrected for the false discovery rate (Genovese *et al*. 2002) for descriptive purposes.

4. Results

4.1. Behavioural Results

Here we start by giving a brief summary of the most important behavioral results

for the whole group reported in Folia *et al*. (2008) and then focus on the specifics for the sub-sample for which CNTNAP2 (SNP RS7794745) data were available. As in previous studies (Forkstam *et al.* 2008), the classification performance of the whole group was well above chance for both instruction types (preference classification: $P < .001$; grammaticality classification: $P < .001$). Standard signal detection analysis showed a robust d-prime effect in discriminating between G and NG sequences (preference: $P < .001$; grammaticality: $P < .001$). No significant response bias was found (preference and grammaticality classification $P > .6$). Participants did not discriminate between high and low ACS sequences (preference: $P > .22$; grammaticality: $P > .66$), and there was no significant response bias (preference $P > .98$; grammaticality: $P > .8$).

We then analyzed the performance data in terms of endorsement rate (i.e. item classified as grammatical independent of their actual grammaticality status). In other words, if the subjects acquire significant aspects of the grammar, then they should endorse grammatical items more often than non-grammatical items. Both grammaticality status and local subsequence familiarity influenced the endorsement rate. The endorsement rate was significantly affected by grammaticality status (preference: $P < .001$; grammaticality: $P < .001$), and by local subsequence familiarity (preference: $P < .001$; grammaticality: $P < .001$), while the interaction between grammaticality status and local subsequence familiarity was non-significant (preference: $P = .06$; grammaticality: $P = .11$). These results show that grammaticality status is used for structural generalization in classifying novel sequences and thus provide support for the notion that grammatical structure instead of subsequence, or fragment features, determine classification (Folia *et al*. 2008).

The critical measure in the behavioral results was the preference of the participants for grammatical, and relative aversion of non-grammatical, sequences. The participants only need to indicate whether they like or dislike a given sequence and therefore we do not need to inform them about the presence of a complex rule system before classification (or at any other point of the experiment), which is the case in standard versions of the AGL paradigm, which uses grammaticality instead of preference classification. Therefore, from the subject's point of view, there is no such thing as a correct or incorrect response and the motivation to use explicit strategies is thus minimized. The participants were also strongly encouraged to trust their gut-feeling in making their decisions. Consistent with this, the subjective reports from the structured post-experimental interview showed that the participants did not utilize an explicit strategy but that their classification decisions were based on gut-feeling. Moreover, the subjective ratings of perceived performance did not correlate with the actual classification performance (Folia *et al*. 2008).

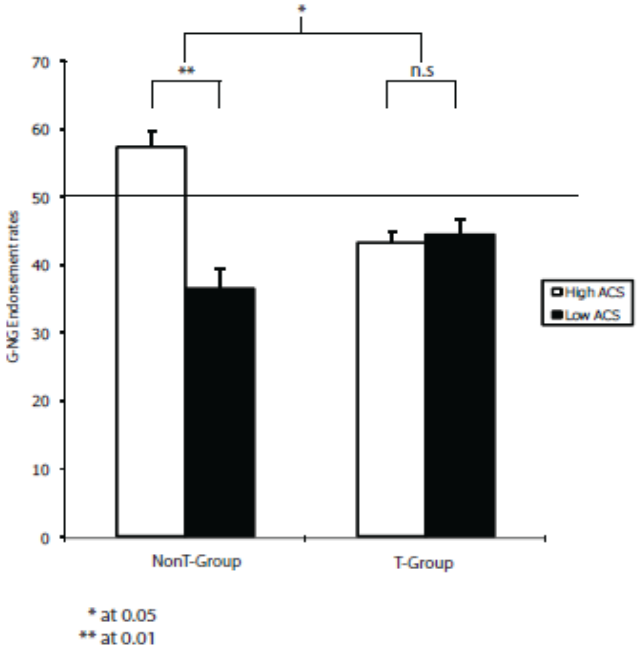

Figure 1: Grammaticality classification and CNTNAP2. The endorsement rates for grammatical and non-grammatical sequences in the T- and nonT-group. The interaction between grammaticality status and local subsequence familiarity was significant for the nonT-group (AA carriers) and not the T-group. The nonT-group thus shows greater dependence on local subsequence familiarity in making the grammaticality judgments than the T-group, despite the fact that local subsequence familiarity is not predictive for grammaticality status. Error bars corresponds to standard error of the mean.

Overall, the sub-sample for which CNTNAP2 data were available was found to behave essentially identical to the whole group and here we focus on their grammaticality classification performance. On the last day, the correct classification performance was well above chance on grammaticality classify-cation (78 ± 19% correct, $T(11) = 5.36$, $P < .001$). Both grammaticality status and local subsequence familiarity influenced the endorsement rate. Repeated measures ANOVA showed significant main effects of grammaticality status ($F(1,11) = 13.2$, $P = .004$) and local subsequence familiarity ($F(1,11) = 21.0$, $P = .001$). We then analyzed the data with a repeated measure ANOVA with gram-maticality status and local subsequence familiarity (ACS) as within-subject variables and allele (T/nonT) as between factors. Post-hoc analysis was con-ducted where relevant. The correct classification performance was significantly greater than chance in both groups (T-group: $T(7) = 3.34$, $P = .01$; nonT-group: $T(3) = 8.25$, $P = .004$). For grammaticality classification, the three-way interaction between grammaticality status, local subsequence familiarity, and allele group was significant ($F(1,10) = 4.86$, $P < .05$) as well as the main effect of grammati-cality status ($F(1,10) = 20.5$, $P = .001$) and local subsequence familiarity ($F(1,10) = 23.4$, $P = .001$). No other interaction reached significance. Post-hoc analysis in the nonT-group revealed a main effect of grammaticality status ($F(1,3) = 17.5$, $P = .02$), and a significant interaction between grammaticality status and local sub-sequence familiarity ($F(1,3) = 22.6$, $P = .01$). In the T-group, a significant main effect was found for both grammaticality status ($F(1,7) = 11.66$, $P = .01$) and local

subsequence familiarity ($F(1,7) = 17.9$, $P = .004$), while no interaction was significant (Figure 1).

These results show that the T- and the nonT-group behave similarly to the whole sample, including the development of a preference for grammaticality (Folia *et al*. 2008, Forkstam *et al*. 2008). However, the grammaticality classification performance of the T-group was independent of local subsequence familiarity (Figure 1), while this was not the case for the nonT-group. Thus, the absence of a T nucleotide in the CNTNAP2 SNP RS7794745 might be associated with a greater reliance on local subsequence familiarity (ACS) during classification. This, despite the fact that the grammaticality status is independent of local subsequence familiarity, by the construction of the stimulus material, and therefore ACS has little, if any, predictive value with respect to grammaticality status.

4.2. fMRI Results

Here we briefly summarize the results reported in Folia *et al*. (2011). Preference classification compared to the sensorimotor decision baseline (Figure 2) activated a set of brain regions (cluster $P_{FWE} < .001$) very similar to what has been observed in previous studies of grammaticality classification (Petersson *et al*. 2004, 2010, Forkstam *et al*. 2006). These activations included the inferior and middle frontal regions bilaterally (BA 44/45), extending into surrounding cortical regions, frontal operculum, and the anterior insula. Additional prefrontal activations included the anterior cingulate and surrounding cortex. Bilateral posterior activations included the inferior parietal cortex (BA 39/40, extending into the posterior superior temporal (BA 22), bilaterally. Bilateral occipital activations were centered on the middle and inferior occipital gyri and extended into the fusiform and the posterior mid-inferior temporal regions, as well as the cerebellum. Significant activations were also observed in the basal ganglia bilaterally, including the caudate nucleus, globus pallidus, and putamen. The results were similar for 'correctly' preferred HG- and LG sequences (Figure 2). Large, and highly significant, deactivations were found in the bilateral medial temporal lobe memory system, including the hippocampus proper (cluster $P_{FWE} < .001$), replicating previous results for grammaticality classification (Petersson *et al*. 2010).

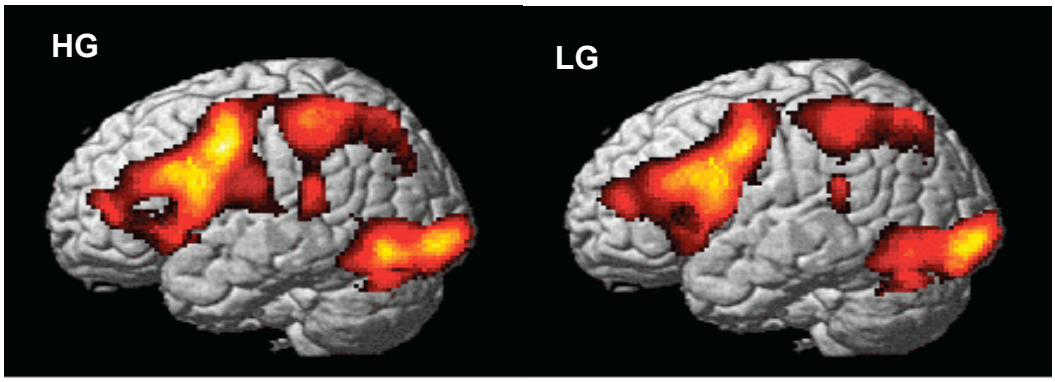

Figure 2: Preference classification. Brain regions engaged during 'correct' preference classification of grammatical sequences with high (HG) and low (LG) subsequence familiarity (ACS) relative the sensorimotor decision baseline. Adapted from Folia et al. (2011).

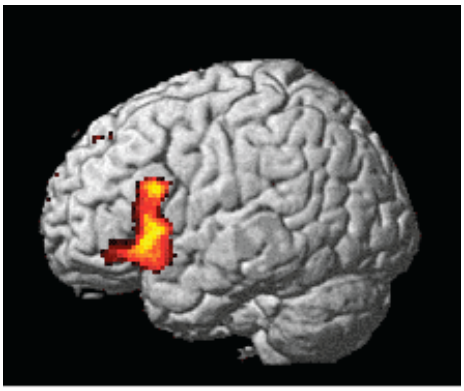

Figure 3: Preference classification. Brain regions engaged by artificial syntactic anomalies (NG > G). Adapted from Folia et al. (2011).

In preference classification (Folia *et al*. 2011), as in previous studies of grammaticality classification (Petersson *et al*. 2004, 2010, Forkstam *et al*. 2006), artificial syntactic anomalies (NG > G; Figure 3) engaged a network of brain regions, including the left inferior and right inferior-middle frontal gyrus (left and right cluster $P_{FWE} < .001$) centered on Broca's region (BA 44/45). In the reverse contrast (G > NG), we observed no significant differences. There was no significant effect of local subsequence familiarity (cluster $P_{FWE} > .98$), neither were there any significant interaction (cluster $P_{FWE} > .83$), consistent with our behavioral findings (Folia *et al*. 2008).

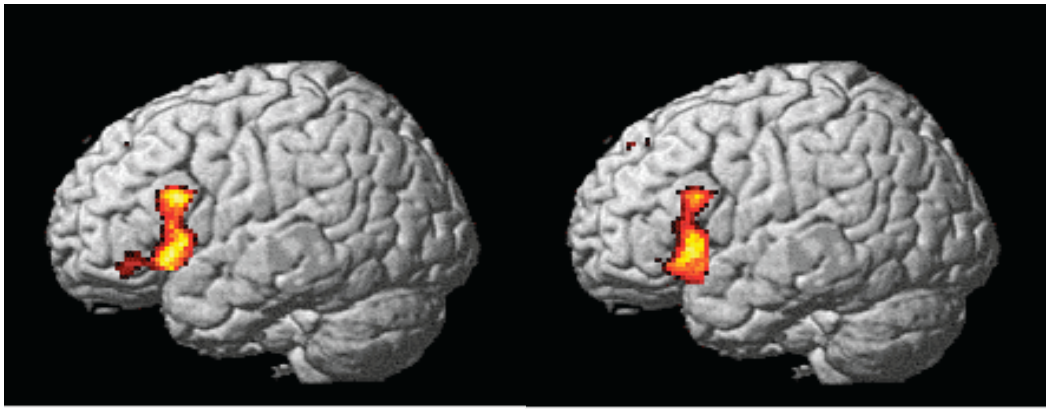

Figure 4: Brain regions engaged during both preference and grammaticality classification. Left: The NG > G effect of Folia et al. (2011) masked with the related effect observed in Petersson et al. (2010). Right: The overlap of the NG > G effect in preference classification (Folia et al. 2011) masked with natural syntax related variability in the same subjects observed in (Folia et al. 2009).

Here, we examined the overlap between preference and grammaticality classification by masking the preference classification contrast (NG vs. G effect) from Folia *et al*. (2011) with the same contrast of grammaticality classification from Petersson *et al*. (2010; Figure 4 and Appendix B). We found a common overlap in the inferior frontal regions, centered on Broca's region (BA 44/45) and extending into the frontal operculum/anterior insula, bilaterally, as well as the right middle frontal region (LIFG cluster $P_{FWE} = .003$; RI/MFG cluster $P_{FWE} < .001$). In addition, the anterior cingulate/supplementary motor regions were found to

be active in both the tasks (ACC/SMA cluster P_{FWE} = .001; see Appendix B for details). Reversing the order of masking yielded identical results (LIFG cluster: P_{FWE} = .003; RI/MFG cluster: P_{FWE} < .001; ACC/SMA cluster P_{FWE} = .001). Moreover, there was no significant difference between preference and grammaticality classification in any contrast, including the main effects of grammaticality status and local subsequence familiarity. Thus, artificial syntax processing engaged the same brain regions during preference and grammaticality classification, although there was a tendency that grammaticality classification yielded somewhat more robust results, highly consistent with the behavioral results (Folia *et al.* 2008, see also Forkstam *et al.* 2008). The same conclusion is reached when we examined the common overlap between artificial and natural syntax processing by masking the NG vs. G effect observed for preference classification with the natural-syntax-related variability in the same subjects (Figure 4; LIFG cluster: P_{FWE} = .001; RI/MFG cluster: P_{FWE} = .008; ACC/SMA cluster P_{FWE} = .001), that is, the main effect of syntax in the 2x2 natural language experiment of Folia *et al*. (2009).

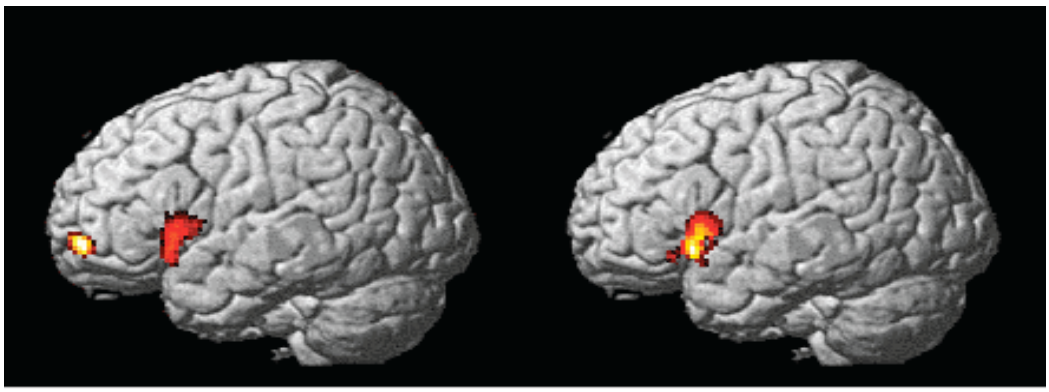

Figure 5: Brain regions differentiating the T- and the nonT-groups. Left: Group differences related to grammaticality classification (nonT > T). Right: Group differences related to grammatical sequences of high local subsequence familiarity (nonT > T).

Finally, we explored the fMRI results of Petersson *et al*. (2010; Figure 4) with respect to differences between the T- and nonT-group. The results showed significantly greater activity for the nonT-group compared to the T-group in the left inferior frontal gyrus (BA 44/45, P_{FWE} = .002), the left fronto-polar region (BA 10, P_{FWE} = .012), and the left ventral occipito-temporal region (BA 37, P_{FWE} = .003) during grammaticality classification. The group difference found in Broca's region was mainly related to differences between the T- and nonT-group when processing grammatical sequences, in particular grammatical sequences of high local subsequence familiarity (BA 44/45 centered on [–48, 16, –2], P_{FWE} = .024; Figure 5). The results were almost identical for the preference classification data of Folia *et al*. (2011).

5. Discussion

One of the main objectives of this study was to compare the brain networks engaged by preference classification and the standard grammaticality classification

task after implicit artificial syntax acquisition. The results show that preference and grammaticality classification engage virtually identical brain regions, consistent with previously reported behavioral findings (Folia *et al.* 2008, Forkstam *et al.* 2008). The theoretical advantage of preference compared to grammaticality classification is that there is no correct or incorrect response from the perspective of the participant and at no point is there a need to inform the participant about the existence of an underlying generative grammar, as is the case of the standard grammaticality classification. Nevertheless, the results show that preference and grammaticality classification are (qualitatively) equivalent both at the behavioral and brain levels. In particular, Broca's region, the left inferior frontal gyrus centred on BA 44/45, is active during the artificial syntax processing of well-formed (grammatical) sequence independent of local subsequence familiarity. Moreover, this region is engaged to a greater extent when a syntactic anomaly is present and the unification of structural treelets becomes difficult or impossible. The behavioral results of Folia *et al*. (2008) show that subjects implicitly acquired significant knowledge from being exposed to only grammatical examples and without receiving performance feedback at any stage of the experiment. Moreover, the behavioral results show that participants apply implicitly acquired structural knowledge (independent of subsequence familiarity) and the corresponding fMRI results show that brain regions central to natural syntax processing are engaged (Folia *et al.* 2011), also when they are not explicitly instructed or receives any information concerning the existence of a generative grammar. The results of this study show that the participants do so at levels comparable to grammaticality classification. Thus, the structural mere-exposure effect is a robust phenomenon at the behavioral (Folia *et al.* 2008, Forkstam *et al.* 2008) and brain level (Folia *et al.* 2011). In other words, the effects related to artificial syntax processing in the left inferior frontal region (BA 44/45) were essentially identical when we masked these with activity related to grammatical classifycation in the same subjects, as well as when masked with activity related to natural syntax processing in the same participants. Our results are also highly consistent with functional localization of natural language syntax in the left inferior frontal gyrus (Bookheimer 2002, Petersson *et al.* 2004, Hagoort 2005).

We used a simple right-linear unification grammar with a finite vocabulary of terminal symbols and a finite lexicon of primitive trees (*treelets*, i.e. structured lexical items; see materials and methods section for details). From an abstract point of view, unification (Vosse & Kempen 2000) is a way to implement computational control in lexicalist grammars (Forkstam & Petersson 2005, Petersson *et al*. 2005). More specifically, for a given lexical item of the grammar used in this study, for example $[s_j, [T, s_k]]$, the features s_j, s_k can be interpreted as control features and T as a surface feature. Here, two lexical items, $[s_i, [R, s_j]]$ and $[s_k, [Q, s_l]]$, unify (i.e. combine or merge) through a *unification operation* U if and only if $s_j = s_k$, or $s_l = s_i$, a process which is incremental and recursive. For example, if the structure $[s_1, [M, [s_2, [S, s_2]]]]$ is already present in the unification space when the lexical item $[s_2, [V, s_4]]$ is retrieved, a larger combinatorial structure can be formed by unification $U([s_1, [M, [s_2, [S, s_2]]]], [s_2, [V, s_4]]) = [s_1, [M, [s_2, [S, [s_2, [V, s_4]]]]]]$, and so on. We note that the control features have acquired a particular functional role in this picture, which can be described in terms of governing the unification

process based on selecting the structural arrangement that can be integrated. In a certain sense therefore, the finite-state control has been distributed over the lexicon among the lexical items in terms of control features. In essence, this retraces a major trend in theoretical linguistics in which more of the grammar is shifted into the lexicon and the distinction between lexical items and grammatical rules is beginning to vanish (cf. Joshi & Schabes 1997, Vosse & Kempen 2000, Jackendoff 2002, 2007). In this context, Broca's region can be considered as a brain region that gradually controls the outcome of parsing or generation. A related, but different proposal has recently been put forward by Bornkessel-Schlesewsky *et al.* (2010), who argue that argue that the left inferior frontal region, including Broca's region, can be described as a brain region that controls the outcome of different processes from general to specific along the anterior-posterior direction. Bornkessel-Schlesewsky *et al.* note that their proposal is partly compatible with Hagoort's (2005) assumption of a unification gradient within the left inferior frontal gyrus.

5.1. A Genetic Basis for Implicit Acquisition of Structured Sequence Knowledge

Two facts about language learning seem indisputable: (i) only humans acquire language, no other species, and thus there must be some biological element that accounts for this ability; (ii) it is also clear that no matter how much of a head start the learner gains through innate constraints, language is learned. Both innate endowment and learning contribute to language acquisition, the result of which is a complex and sophisticated body of linguistic knowledge (Chomsky 1963, Chomsky & Miller 1963). It is clear that unless restrictions are placed on the available "space of possible languages" (i.e. the model space) and/or the characteristics of the acquisition mechanism (i.e. the learning dynamics), "learning" would simply reduce to storing experience (Petersson 2005a, Folia *et al.* 2010). Much of the current discussion of language acquisition concerning the nature of innate constraints is focused on whether these are linguistically specific or not (e.g. Chomsky 1986, 2005; however, see Nowak *et al*. 2002, Chomsky 2007, Christiansen & Chater 2008, Hornstein 2009). We think this is an empirical issue — however, what is clear is that no interesting, complex form of learning is possible without constraints (Vapnik 1998, Jain *et al*. 1999). In this context, Yang (2004) cites an interesting insight by Jerry Fodor (2001: 107–108), "Chomsky can with perfect coherence claim that innate, domain specific [constraints] mediate language acquisition, while remaining entirely agnostic about the domain specificity of language acquisition mechanisms". What can this possibly mean? Folia *et al*. (2010) outline several possibilities. For instance, the learning/developmental dynamics might be domain-general in form, but in the context of language acquisition, operate on a model space that is restricted by innate, language-specific constraints. By language-specific constraints we mean constraints which play no role in cognition outside the language faculty. No one doubts the existence of innate constraints, rather the issue is whether the innate constraints are specific to language or not. In fact, Folia *et al.* argue that in order to rule out innate, language-specific constraints completely, it is necessary to establish that none of the following candidates carry such constraints: (1) the initial state of the

learner; (2) the model space; (3) the learning/developmental dynamics; (4) the representational space; or (5) the representational dynamics — a difficult empirical task. Alternatively, if sufficient non-language-specific constraints for language acquisition are discovered, the necessity of language-specific constraints recedes.

In this fMRI study we took advantage of the fact that a subsample of our participants (Petersson *et al.* 2010, Folia *et al.* 2011) was part of the BIG project at the Donders Centre for Cognitive Neuroimaging and the Department of Human Genetics of the Radboud University Nijmegen. This allowed us to explore the potential role of the CNTNAP2 gene in artificial syntax acquisition at both the behavioural and the brain level. This small scale investigation of possible CNTNAP2 related effects (more precisely, effects related to the common polymorphism observed at the single nucleotide polymorphism RS7794745) in the context of artificial syntax acquisition and structured sequence processing suggests that the T-group (AT- and TT-carriers) was sensitive to the grammaticality status of the sequences independent of local subsequence familiarity. This might mean that individuals with this genotype acquire structural knowledge more rapidly, utilize the acquired knowledge more effectively, or are better able to ignore cues related to local subsequence familiarity in comparison to the nonT-group (AA carriers). This suggests differences in the implicit acquisition process between the two groups. Another possibility is that, if the two groups eventually achieve the same level of successful overall classification at the end acquisition, the nature of sequence processing might be different, since only the nonT-group is sensitive to local subsequence familiarity (which is not predictive of the grammaticality status). In contrast, the T-group relies only (or at least to a greater extent) on their implicitly acquired structural knowledge, which they successfully generalize to novel items. This suggests a qualitative, rather than a quantitative, processing difference between groups. Parallel to these behavioral findings, we observed significantly greater activation in Broca's region centered on the left BA 44/45 as well as the left frontopolar region (BA 10) in the nonT- compared to the T-group. The meaning of these fMRI differences between the two groups is unclear and requires further research for a full understanding. Nevertheless, these initial efforts suggest that it is worthwhile to investigate the genetic basis of the capacity for structured sequence processing in large-scale studies by investigating the relevant biological pathway(s) (Konopka & Geschwind 2010, Newbury & Monaco 2010, Pezawas & Meyer-Lindenberg 2010, and references therein). However, given that CNTNAP2 has been linked to specific language impairment (SLI) and provides a mechanistic link between clinically distinct syndromes involving disrupted language (Vernes *et al.* 2008), and assuming that the structured sequence learning mechanism investigated by artificial grammar learning is shared between artificial and natural syntax acquisition, the present behavioral and fMRI results might suggest that the FOXP2–CNTNAP2 pathway is somehow related to the acquisition of structured sequence knowledge as well as individual differences in artificial and natural syntax acquisition.

5.2. Language as a Neurobiological System and Bounded Recursion

Cognitive neuroscience approaches the brain as a computational system — a sys-

tem conceptualized in terms of information processing. This entails the idea that a subclass of its physical states is viewed as representations and that transitions between states can be understood as a process implementing operations on the corresponding representational structures. It is uncontroversial that any physically realizable computational system is necessarily finite with respect to its memory organization and that it processes information with finite precision (e.g., due to the presence of internal noise or architectural imprecision; Turing 1936a, 1936b, Minsky 1967, Savage 1998, Koch 1999). We have previously indicated why this state of affairs renders the Chomsky hierarchy for classical cognitive models (i.e. Church–Turing computational models) less relevant to neurobiological systems from a neurobiological processing perspective (Petersson 2005a, 2005b, 2008, Petersson *et al.* 2010). The Chomsky hierarchy is in essence a memory hierarchy and it distinguishes between (a few) complexity classes (and corresponding grammar classes) in the context of infinite (unbounded) memory. If we view the faculty of language as a neurobiological system, given its finite storage capacity and finite precision computation, the Chomsky hierarchy is less relevant — it does not make the relevant distinctions. However, bounded versions of the different memory architectures entailed by the hierarchy might be relevant (although we think these should not be taken too seriously). For example, the unbound push-down stack is a memory architecture corresponding to the class of context-free grammars, and it is conceivable that a bounded push-down stack is used in language processing, as suggested by Levelt (1974) as one possibility. Of course, this does not imply that the Chomsky hierarchy is irrelevant for computational theory (Davis *et al.* 1994, Pullum & Scholz 2010) or competence grammars in theoretical linguistics (Chomsky 1963). However, we note that modern complexity theory, which is more closely related to processing complexity rather than the Chomsky hierarchy, makes fine-grained distinctions (Cutland 1980, Papadimitriou 1993, Savage 1998, Hopcroft *et al*. 2000, Arora & Barak 2009) and might, perhaps, be useful from a neurobiological processing perspective (although this is unclear).

With the advent of generative grammar, recursion became key to achieving discrete infinity (e.g. Chomsky 1956, 1963). Accordingly, early psycholinguistics devoted considerable effort to the study of complex recursive constructions, especially in the form context-free or more general grammars (Chomsky 1963, Levelt 1974). However, it was theoretically suggested (e.g. Chomsky 1963: 329–333, 390), and soon empirically confirmed, that unbound (i.e. infinite) recursive capacity is not realizable in human performance (~actual cognitive processing). Thus, it was found that sentences with more than two center-embeddings are read with the same intonation as a list of random words (Miller 1962), cannot easily be memorized (Miller & Isard 1964, Foss & Cairns 1970), are difficult to paraphrase (Hakes & Cairns 1970, Larkin & Burns 1977) and comprehend (Wang 1970, Hamilton & Deese 1971, Blaubergs & Braine 1974, Hakes *et al*. 1976), and are, paradoxically, judged to be ungrammatical (Marks 1968).

Recursion is once again attracting attention as an hypothesized key feature of the language faculty, with the suggestion that unbounded recursion may be the only property of the language faculty that is both species-specific and domain-specific (Hauser *et al.* 2002). Nevertheless, in order to preserve the essen-

tial feature of the notion of discrete infinity (unbounded "human creativity"), Chomsky introduced the notion of a competence grammar, "a device that enumerates […] an infinite class of sentences with structural descriptions" (Chomsky 1963: 329–330, device A in Fig. 1). The competence grammar is distinct from the language acquisition and processing ("performance") system (Chomsky 1963: 329–330, devices C and B, respectively, in Fig. 1). One consequence of grammars or computational models that support unbounded recursion (and infinite precision processing), is that they overgeneralize, by generating arbitrarily long sequences (and correspondingly complex sequence structures) that are never used, and in fact, has never been observed. This might or might not be a problem, depending on ones perspective on these issues. However, this is not a problem for bounded recursive procedures (or equivalent analogues, Petersson 2005b, 2008). As previously noted, one uncontroversial limitation on actual neurobiological systems is their finiteness, both in terms of memory and processing precision. For instance, Chomsky remarks that both language processing and language acquisition, "which represents actual performance, must necessarily be strictly finite", that is, a finite-state machine (Chomsky 1963: 331–333); and continues: "Nevertheless, the performance of the speaker or hearer must be representable by a finite automaton of some sort" (p. 390). However, he further argued that "any interesting realization of B [i.e. a finite-state processing system] that is not completely ad hoc will incorporate A [i.e. a competence grammar] as a fundamental component". One example of this idea is a (e.g., universal) Turing machine with finite tape-memory (Petersson *et al.* 2010: fn. 3). Another example is a (e.g., universal) register machine with a finite number of registers (Petersson 2005b). In both cases, it could be argued that the finite-state control unit, in a certain sense, represents unbounded 'knowledge' (or competence grammar) as well as unbounded recursive potential. However, this knowledge cannot be fully expressed, and the recursive potential not fully realized, because of memory limitations. But it could be argued, as Chomsky (1963) does, that if we imagine that hardware constraints can be disregarded (abstracted away), then the system instantiates the equivalent of a competence grammar, and thus unbounded 'knowledge', in this sense. Perhaps one way to interpret this idea, when applied to the language faculty, is in analogy with frictionless mechanics in physics — it retains instrumental value, but is not a correct description of the underlying reality (e.g., a correct model of friction is an atomic, mainly electromagnetic phenomenon).

Finite-state and finite-precision computation devices, including real neural networks, are sufficient to handle bounded recursion of general type, so there is no real problem here from the point of view of language processing ('performance'). We think this opens the possibility for lateral thinking on matters related to the knowledge of language ('competence'). We argue that more realistic neural models provide natural bounds on memory and on processing as well as architectural precision, and therefore, on the specification of the language faculty viewed as a neurobiological system (cf. Petersson *et al.* 2010). Generally, analog dynamical systems provide a non-classical information processing alternative to classical computational architectures (Siegelmann & Fishman 1998). In particular, network approaches offer possibilities to model cognition within a

non-classical dynamical systems framework that is natural from a neurobiological perspective. It is known theoretically, that under the assumption of infinite precision processing, Church-Turing computable processes can be embedded in dynamical systems instantiated by neural networks (e.g. Siegelmann 1999). For example, the discrete-time recurrent network can be viewed as a simple network analogue of the finite-state architecture (Petersson 2005b, Petersson *et al.* 2005). In general, the recurrent neural network architecture can be viewed as an architecture with a finite number of dynamic, analog registers (e.g., the "membrane potential") that processes information interactively. In the simplest case, computations are determined by the network topology and by the transfer functions of the processing units, as well as the set of dynamical variables associated with these processing units. Moreover, important aspects of both short-term and long-term memory are co-localized with processing infrastructure (Petersson 2005a, Petersson *et al*. 2009). From a neurobiological perspective, therefore, it seems natural to try to understand language acquisition and language processing in terms of adaptive dynamical systems (Petersson 2005a, Petersson *et al.* 2009, 2010). Thus, an important challenge in the neurobiology of syntax is to understand syntax processing in terms of noisy spiking network processors. Similar, independent, accounts have been put forward by Culicover & Nowak (2003) in their Dynamical Grammar as well as others (Christiansen & Chater 1999, Rodriguez *et al*. 1999, Rodriguez 2001,).

What are the implications of this for theoretical models of language and grammar? The Chomsky hierarchy only has theoretical meaning in the context of infinite memory resources. Rather than giving unbounded recursion the centre stage, some of the important issues in the neurobiology of syntax, and language more generally, are related to the nature of the neural code (i.e. representation), the character of human on-line processing memory, and noisy neural finite precision computation (Koch 1999, Trappenberg 2010). Recurrent connectivity is a generic feature of brain network topology (Nieuwenhuys *et al*. 1988). Thus, recursive processing is a latent capacity in almost any neurobiological system and it would be surprising, indeed, if this feature would be unique to the faculty of language. We noted that one relevant issue from the point of view of natural language is the human capacity to process patterns of non-adjacent dependencies — not arbitrarily 'long' non-adjacent dependencies — there is a definite natural upper-bound set by the brain and its underlying neurophysiology. We can thus choose to work with any fruitful formal syntax framework as long as this serves its purpose, for example, to capture the presence of bounded relational patterns between lexical items in compositionally constructed sentences, to elaborate parameterized model of language acquisition or, if we are not interested in hardware constraints and implementation issues, abstract away the implementation level and explore 'frictionless' models of the language faculty.

6. Conclusion

One of the objectives of this study was to compare the brain networks engaged by artificial syntax processing during preference and grammaticality classifi-

cation after implicit artificial syntax acquisition. The results show that preference and grammaticality classification engage virtually identical brain regions, consistent with previously reported behavioral findings. In particular, the left inferior frontal region centered on BA 44/45 (Broca's region) is active during artificial syntax processing of well-formed sequences independent of local subsequence familiarity. The effects related to artificial syntax in the left inferior frontal region (BA 44/45) were essentially identical when masked with activity related to natural syntax obtained in the same subjects. Thus, the current fMRI results show that artificial syntax processing engages brain regions central to natural syntax processing. We suggest, therefore, that the left inferior frontal region is a generic on-line sequence processor that unifies information from various sources in an incremental and recursive manner. Finally, we explored CNTNAP2 related effects in artificial syntax acquisition and structured sequence processing. The results suggest that AT- and TT-carriers (at the CNTNAP2 SNP RS7794745) were sensitive to the grammaticality status independent of local subsequence familiarity, while AA-carriers were sensitive to local subsequence familiarity. We observed significantly greater activation in Broca's region and the left frontopolar region (BA 10) in the AA-carriers compared to AT- and TT-carriers. The meaning of these behavioural and fMRI findings is unclear and requires further investigation. Nevertheless, these initial efforts suggest that it is worthwhile to try to understand the genetic basis for language as well as the capacity for structured sequence processing in large-scale studies by investigating the relevant biological pathway(s).

Appendix A: Example stimuli used for preference and grammaticality classification

Stimulus Categories	Classification Items
High Grammatical (HG)	VXVRXSVS
	MSSSVRXSV
	VXSSVRXVRXSV
	MVRXSSSSVS
Low Grammatical (LG)	VXSVS
	MSSSSSV
	VXSVRXRRRR
	MSSSVRXRRRRM
High Non-Grammatical (HNG)	VRVRXSSS
	MRXSSSV
	VRXRXSVRXRM
	MVXSVRXVRXRM
Low Non-Grammatical (LNG)	VRXRXRM
	VXVRXVXRM
	MSVRXSXRRM
	MSSVRSSSVS

Appendix B: Overlap between preference and grammaticality classification

Anatomical region	*Brodmann's area*	[*x y z*]	*Z-score*	*P-value*
Left Inferior Frontal Cluster				***.003***
L inferior frontal gyrus	BA 44	–54 14 2	4.07	.013
	BA 44/45	–60 20 16	3.90	.016
	BA 44/45	–52 18 22	3.81	.019
	BA 45	–60 20 10	3.54	.028
	BA 45	–46 22 22	3.35	.039
	BA 45/47	–56 18 2	4.02	.014
	BA 47	–42 20 –10	3.90	.016
L frontal operculum/anterior insula	BA 49/13/15	–38 18 –10	4.02	.010
Right Inferior-Middle Frontal Cluster				***< .001***
R inferior frontal gyrus	BA 44/45	50 24 18	3.65	.023
	BA 45	56 30 12	3.54	.028
	BA 45/47	58 32 0	3.42	.035
	BA 47	46 32 –4	5.08	.010
	BA 47/11	48 44 –14	3.58	.026
R mid-anterior insula	BA 13/15	40 20 –6	4.15	.011
R frontal operculum/anterior insula	BA 49/15	36 20 –10	3.87	.017
	BA 49/13/15	32 26 0	3.57	.027
R inferior-middle frontal gyrus	BA 45/46	46 34 12	3.41	.036
	BA 45/46	48 30 16	3.39	.037
	BA 45/46	58 34 16	3.37	.038
	BA 45/46	58 34 8	3.23	.047
	BA 46	52 40 18	3.27	.044
Medial Prefrontal-Frontopolar Cluster				***.001***
Anterior cingulate/supplementary motor	BA 8	0 26 52	4.41	.010
	BA 8/32	6 30 44	4.60	.010
	BA 8/32	14 16 48	3.38	.037
	BA 6/8/32	–6 14 54	4.04	.013
Anterior cingulate cortex	BA 32	8 34 38	4.53	.010
	BA 32	10 32 24	4.27	.010

Local maxima observed for correctly classified non-grammatical vs. grammatical items. Cluster P-values are family-wise error corrected and P-values of local maxima are corrected based on the false-discovery rate.

References

Abrahams, Brett S., Dmitri Tentler, Julia V. Perederiy, Michael C. Oldham, Giovanni Coppola & Daniel H. Geschwind. 2007. Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17849–17854.

Adler, Robert J. 1981. *The Geometry of Random Fields*. New York: Wiley and Sons.

Adler, Robert J. & Jonathan E. Taylor. 2007. *Random Fields and Geometry*. New York: Springer.

Alarcón, Maricela, Brett S. Abrahams, Jennifer L. Stone, Jacqueline A. Duvall, Julia V. Perederiy, Jamee M. Bomar, Jonathan Sebat, Michael Wigler, Christa L. Martin, David H. Ledbetter, Stanley F. Nelson, Rita M. Cantor & Daniel H. Geschwind. 2008. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *American Journal of Human Genetics* 82, 150–159.

Alberts, Bruce, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts & Peter Walter. 2007. *Molecular Biology of the Cell*. New York: Garland Science.

Arora, Sanjeev & Boaz Barak. 2009. *Computational Complexity: A Modern Approach*. Cambridge: Cambridge University Press.

Bishop, Dorothy. 2009. Genes, cognition, and communication. *Annals of the New York Academy of Sciences* 1156, 1–18.

Blaubergs, Maija S. & Martin D. S. Braine. 1974. Short-term memory limitations on decoding self–embedded sentences. *Journal of Experimental Psychology* 102, 745–748.

Bookheimer, Susan. 2002. Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25, 151–188.

Bornkessel-Schlesewsky, Ina, Tanja Grewe & Matthias Schlesewsky. 2010. Prominence vs. aboutness in sequencing: A functional distinction within the left inferior frontal gyrus. *Brain and Language* doi:10.1016/j.bandl.2010.06.004.

Chomsky, Noam. 1956. Three models for the description of language. *IEEE Transactions on Information Theory* 2, 113–124.

Chomsky, Noam. 1963. Formal properties of grammars. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 2, 323–418. New York: John Wiley.

Chomsky, Noam. 1986. *Knowledge of Language*. New York: Praeger.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam. 2007. Of mind and language. *Biolinguistics* 1, 9–27.

Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 2, 269–321. New York: John Wiley.

Christiansen, Morten H. & Nick Chater. 1999. Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science* 23, 157–205.

Christiansen, Morten H. & Nick Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31, 489–558.

Christiansen, Morten H., Louise M. Kelly, Richard C. Shillcock & Katie Greenfield. 2010. Impaired artificial grammar learning in agrammatism. *Cognition* 116, 382–393.

Conway, Christopher N. & David B. Pisoni. 2008. Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences* 1145, 113–131.

Culicover, Peter W. & Andrzej Nowak. 2003. *Dynamical Grammar: Volume Two of Foundations of Syntax*. Oxford: Oxford University Press.

Cutland, Nigel J. 1980. *Computability: An Introduction to Recursive Function Theory*. Cambridge: Cambridge University Press.

Davidson, Eric H. 2006. *The Regulatory Genom: Gene Regulatory Networks in Development and Evolution*. San Diego, CA: Academic Press.

Davidson, Eric H., Jonathan P. Rast, Paola Oliveri, Andrew Ransick, Cristina Calestani, Chiou-Hwa Yuh, Takuya Minokawa, Gabriele Amore, Veronica Hinman, Cesar Arenas-Mena, Ochan Otim, C. Titus Brown, Carolina B. Livi, Pei Yun Lee, Roger Revilla, Alistair G. Rust, Zheng jun Pan, Maria J. Schilstra, Peter J. C. Clarke, Maria I. Arnone, Lee Rowen, R. Andrew Cameron, David R. McClay, Leroy Hood & Hamid Bolouri. 2002. A genomic regulatory network for development. *Science* 295, 1669–1678.

Davis, Martin, Ron Sigal & Elaine J. Weyuker. 1994. *Computability, Complexity, and Languages: Fundamentals of Theoretical Computer Science*. San Diego, CA: Academic Press.

Enard, Wolfgang, Molly Przeworski, Simon E. Fisher, Cecilia S. L. Lai, Victor Wiebe, Takashi Kitano, Anthony P. Monaco & Svante Pääbo. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418, 869–872.

Fitch, W. Tecumseh & Marc D. Hauser. 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 303, 377–380.

Fodor, Jerry A. 2001. Doing without *What's Within*: Fiona Cowie's criticism of nativism. *Mind* 110, 99–148.

Folia, Vasiliki, Christian Forkstam, Peter Hagoort & Karl Magnus Petersson. 2009. Language comprehension: The interplay between form and content. *Proceedings of the Cognitive Science Society* 2009, 1686–1691.

Folia, Vasiliki, Christian Forkstam & Karl Magnus Petersson. 2011. Implicit structured sequence learning. Submitted.

Folia, Vasiliki, Julia Uddén, Christian Forkstam, Martin Ingvar, Peter Hagoort & Karl Magnus Petersson. 2008. Implicit learning and dyslexia. *Annals of the New York Academy of Sciences* 1145, 132–150.

Folia, Vasiliki, Julia Uddén, Christian Forkstam & Karl Magnus Petersson. 2010. Artificial language learning in adults and children. *Language Learning* 60(s2), 188–220.

Forkstam, Christian, Åsa Elwér, Martin Ingvar & Karl Magnus Petersson. 2008. Instruction effects in implicit artificial grammar learning: A preference for grammaticality. *Brain Research* 1221, 80–92.

Forkstam, Christian, Peter Hagoort, Guillén Fernández, Martin Ingvar & Karl Magnus Petersson. 2006. Neural correlates of artificial syntactic structure classification. *NeuroImage* 32, 956–967.

Forkstam, Christian & Karl Magnus Petersson. 2005. Syntactic classification of acquired structural regularities. *Proceedings of the Cognitive Science Society* 2005, 696–701.

Foss, Donald J. & Helen S. Cairns. 1970. Some effects of memory limitations upon sentence comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 9, 541–547.

Friston, Karl J., John T. Ashburner, Stefan J. Kiebel, Thomas E. Nichols, William D. Penny (eds.). 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. San Diego, CA: Academic Press.

Genovese, Christopher R., Nicole A. Lazar & Thomas Nichols. 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15, 870–878.

Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207.

Gómez, Rebecca L. & LouAnn Gerken. 2000. Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences* 4, 178–186.

Hagoort, Peter. 2005. On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences* 9, 416–423.

Hakes, David T., Judith S. Evans & Linda L. Brannon. 1976. Understanding sentences with relative clauses. *Memory & Cognition* 4, 283–290.

Hakes, David T. & Helen S. Cairns. 1970. Sentence comprehension and relative pronouns. *Perception & Psychophysics* 8, 5–8.

Hamilton, Helen W. & James Deese. 1971. Comprehensibility and subject-verb relations in complex sentences. *Journal of Verbal Learning & Verbal Behavior* 10, 163–170.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Hockett, Charles F. 1963. The problem of universals in language. In Joseph H. Greenberg (ed.), *Universals of Language*, 1–29. Cambridge, MA: MIT Press.

Hockett, Charles F. 1987. *Refurbishing Our Foundations: Elementary Linguistics from an Advanced Point of View*. Philadelphia, PA: John Benjamins.

Hopcroft, John E., Rajeev Motwani & Jeffrey D. Ullman. 2000. *Introduction to Automata Theory, Languages, and Computation*. Reading, MA: Addison-Wesley.

Hornstein, Norbert. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, Ray. 2007. A Parallel Architecture perspective on language processing. *Brain Research* 1146, 2–22.

Jain, Sanjay, Daniel Osherson, James S. Royer & Arun Sharma. 1999. *Systems That Learn*. Cambridge, MA: MIT Press.

Joshi, Aravind K. & Yves Schabes. 1997. Tree-adjoining grammars. In Grzegorz Rozenberg & Arto Salomaa (eds.), *Handbook of Formal Languages, vol. 3: Beyond Words*, 69–124. Berlin: Springer Verlag.

Knowlton, Barbara J. & Larry R. Squire. 1996. Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22, 169–181.

Koch, Christof. 1999. *Biophysics of Computation: Information Processing in Single Neurons*. New York: Oxford University Press.

Konopka, Genevieve, Jamee M. Bomar, Kellen Winden, Giovanni Coppola,

Zophonias O. Jonsson, Fuying Gao, Sophia Peng, Todd M. Preuss, James A. Wohlschlegel & Daniel H. Geschwind. 2009. Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature* 462, 213–217.

Konopka, Genevieve & Daniel H. Geschwind. 2010. Human brain evolution: Harnessing the genomics (r)evolution to link genes, cognition, and behavior. *Neuron* 68, 231–244.

Lai, Cecilia S. L., Simon E. Fisher, Jane A. Hurst, Faraneh Vargha-Khadem & Anthony P. Monaco. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523.

Larkin, Willard & David Burns. 1977. Sentence comprehension and memory for embedded structure. *Memory & Cognition* 5, 17–22.

Levelt, Willem J.M. 1974. *Formal Grammars in Linguistics and Psycholinguistics*, vol. III: *Psycholinguistic Applications*. The Hague: Mouton.

MacDermot, Kay D., Elena Bonora, Nuala Sykes, Anne-Marie Coupe, Cecilia S.L. Lai, Sonja C. Vernes, Faraneh Vargha-Khadem, Fiona McKenzie, Robert L. Smith, Anthony P. Monaco & Simon E. Fisher. 2005. Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *The American Journal of Human Genetics* 76, 1074–1080.

Manza, Louis & Robert F. Bornstein. 1995. Affective discrimination and the implicit learning process. *Consciousness & Cognition* 4, 399–409.

Marks, Lawrence E. 1968. Scaling of grammaticalness of self-embedded English sentences. *Journal of Verbal Learning and Verbal Behavior* 7, 965–967.

Meulemans, Thierry & Martial van der Linden. 1997. Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23, 1007–1028.

Miller, George A. 1962. Some psychological studies of grammar. *American Psychologist* 17, 748–762.

Miller, George A. & Stephen Isard. 1964. Free recall of self-embedded English sentences. *Information and Control* 7, 292–303.

Minsky, Marvin L. 1967. *Computation: Finite and Infinite Machines*. Englewood Cliffs, NJ: Prentice-Hall.

Misyak, Jennifer B., Morten H. Christiansen & J. Bruce Tomblin. 2009. Statistical learning of nonadjacencies predicts on-line processing of long-distance dependencies in natural language. *Proceedings of the Cognitive Science Society* 2009, 177–182.

Misyak, Jennifer B., Morten H. Christiansen & J. Bruce Tomblin. 2010a. On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology,* doi:10.3389/fpsyg.2010.00031.

Misyak, Jennifer B., Morten H. Christiansen & J. Bruce Tomblin. 2010b. Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science* 2, 138–153.

Newbury, Dianne F. & Anthony P. Monaco. 2010. Genetic advances in the study of speech and language disorders. *Neuron* 68, 309–320.

Nieuwenhuys, Rudolf, Jan Voogd, Christiaan van Huijzen. 1988. *The Human Central Nervous System: A Synopsis and Atlas,* 3rd revised edn. Berlin: Springer Verlag.

Nowak, Martin A., Natalia L. Komarova & Partha Niyogi. 2002. Computational and evolutionary aspects of language. *Nature* 417, 611–617.

O'Donnell, Timothy J., Marc D. Hauser & W. Tecumseh Fitch. 2005. Using mathematical models of language experimentally. *Trends in Cognitive Sciences* 9, 284–289.

Packard, Mark G., & Barbara J. Knowlton. 2002. Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience* 25, 563–593.

Papadimitriou, Christos H. 1993. *Computational Complexity*. Upper Saddle River, NJ: Addison Wesley.

Petersson, Karl Magnus. 2005a. *Learning and Memory in the Human Brain*. Stockholm: Karolinska University Press.

Petersson, Karl Magnus. 2005b. On the relevance of the neurobiological analogue of the finite-state architecture. *Neurocomputing* 65–66, 825–832.

Petersson, Karl Magnus. 2008. On cognition, structured sequence processing and adaptive dynamical systems. *Proceedings of the American Institute of Physics: Mathematical and Statistical Physics Subseries* 1060, 195–200.

Petersson, Karl Magnus, Vasiliki Folia & Peter Hagoort. 2010. What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, doi:10.1016/j.bandl.2010.08.003.

Petersson, Karl Magnus, Christian Forkstam & Martin Ingvar. 2004. Artificial syntactic violations activate Broca's region. *Cognitive Science* 28, 383–407.

Petersson, Karl Magnus, Peter Grenholm & Christian Forkstam. 2005. Artificial grammar learning and neural networks. *Proceedings of the Cognitive Science Society* 2005, 1726–1731.

Petersson, Karl Mangus, Martin Ingvar & Alexandra Reis. 2009. Language and literacy from a cognitive neuroscience perspective. In David R. Olson & Nancy Torrance (eds.), *Cambridge Handbook of Literacy*, 152–182. Cambridge: Cambridge University Press.

Pezawas, Lukas & Andreas Meyer-Lindenberg. 2010. Imaging genetics: Progressing by leaps and bounds. *NeuroImage* 53, 801–803.

Poliak, Sebastian, Leora Gollan, Ricardo Martinez, Andrew Custer, Steven Einheber, James L. Salzer, James S. Trimmer, Peter Shrager & Elior Peles. 1999. Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K+ channels. *Neuron* 24, 1037–1047.

Pullum, Geoffrey K. & Barbara C. Scholz. 2010. Recursion and the infinitude claim. In Harry van der Hulst (ed.), *Recursion in Human Language*, 113–138. Berlin: Mouton de Gruyter.

Reber, Arthur S. 1967. Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 5, 855–863.

Rodriguez, Paul. 2001. Simple recurrent networks learn context-free and context-sensitive languages by counting. *Neural Computation* 13, 2093–2118.

Rodriguez, Paul, Janet Wiles & Jeffrey L. Elman. 1999. A recurrent neural network that learns to count. *Connection Science* 11, 5–40.

Saffran, Jenny, Marc Hauser, Rebecca Seibel, Joshua Kapfhamer, Fritz Tsao & Fiery Cushman. 2008. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* 107, 479–500.

Savage, John E. 1998. *Models of Computation*. Reading, MA: Addison-Wesley.
Siegelmann, Hava T. 1999. *Neural Networks and Analog Computation: Beyond the Turing Limit*. Basel: Birkhäuser.
Siegelmann, Hava T. & Shmuel Fishman. 1998. Analog computation with dynamical systems. *Physica D* 120, 214–235.
Snijders, Tom M., Mark Rijpkema, Barbara Franke, Han G. Brunner, Dan Dediu, Vasiliki Folia, Julia Uddén, Guillén Fernández, Karl Magnus Petersson & Peter Hagoort. 2011. A common CNTNAP2 polymorphism affects functional and structural aspects of language-relevant neuronal infrastructure. Submitted.
Stadler, Michael A. & Peter A. Frensch (eds.). 1998. *Handbook of Implicit Learning*. London: SAGE.
Talairach, Jean & Pierre Tournoux. 1988. *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System — An Approach to Cerebral Imaging*. New York: Thieme Medical Publishers.
Trappenberg, Thomas P. 2010. *Fundamentals of Computational Neuroscience* (2nd edn.). Oxford: Oxford University Press.
Turing, Alan. 1936a. On computable numbers with an application to the Entscheidungs problem (part 1). *Proceedings of the London Mathematical Society* 42, 230–240.
Turing, Alan. 1936b. On computable numbers with an application to the Entscheidungs problem (part 2). *Proceedings of the London Mathematical Society* 42, 241–265.
Uddén, Julia, Susana Araujo, Christian Forkstam, Martin Ingvar, Peter Hagoort & Karl Magnus Petersson. 2009. A matter of time: Implicit acquisition of recursive sequence structures. *Proceedings of the Cognitive Science Society* 2009, 2444–2449.
Uddén, Julia, Vasiliki Folia, Christian Forkstam, Martin Ingvar, Guillén Fernández, Sebastiaan Overeem, Gijs van Elswijk, Peter Hagoort & Karl Magnus Petersson. 2008. The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain Research* 1224, 69–78.
Uddén, Julia, Martin Ingvar, Peter Hagoort & Karl Magnus Petersson, K. 2011. Broca's region: A causal role in implicit processing of grammars with adjacent and non-adjacent dependencies. Submitted.
Ullman, Michael T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92, 231–270.
Vapnik, Vladimir. 1998. *Statistical Learning Theory*. New York: Wiley and Sons.
Vargha-Khadem, Faraneh, David G. Gadian, Andrew Copp & Mortimer Mishkin. 2005. FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience* 6, 131–138.
Vernes, Sonja C., Dianne F. Newbury, Brett S. Abrahams, Laura Winchester, Jérôme Nicod, Matthias Groszer, Maricela Alarcón, Peter L. Oliver, Kay E. Davies, Daniel H. Geschwind, Anthony P. Monaco & Simon E. Fisher.2008. A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine* 359, 2337–2345.
Vosse, Theo & Gerard Kempen. 2000. Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a

lexicalist grammar. *Cognition* 75, 105–143.
de Vries, Meinou H., Andre R. C. Barth, Sandra Maiworm, Stefan Knecht, Pienie Zwitserlood & Agnes Flöel. 2010. Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *Journal of Cognitive Neuroscience* 22, 2427–2436.
Wang, Marilyn D. 1970. The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior* 9, 398–404.
Watkins, Kate E., Nina F. Dronkers, Faraneh Vargha-Khadem. 2002. Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain* 125, 452–464.
Worsley, Keith J., Sean Marrett, Peter Neelin, Alain C. Vandal, Karl J. Friston & Alan C. Evans. 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping* 4, 58–73.
Yang, Charles D. 2004. Universal grammar, statistics or both? *Trends in Cognitive Sciences* 8, 451–456.
Zajonc, Robert B. 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement* 9, Part 2.
Zizak, Diane M. & Arthur S. Reber. 2004. Implicit preferences: The role(s) of familiarity in the structural mere exposure effect. *Consciousness & Cognition* 13, 336–362.

Vasiliki Folia[1, 2, 3]
vasiliki.folia@fcdonders.nl

Christian Forkstam[4]
christian@forkstam.se

Martin Ingvar[3]
martin.ingvar@ki.se

Peter Hagoort[1, 2]
peter.hagoort@mpi.nl

Karl Magnus Petersson[1, 2, 4]
karl-magnus.petersson@mpi.nl

[1] *Max Planck Institute for Psycholinguistics*
P.O. Box 310
6500 AH Nijmegen
The Netherlands

[2] *Radboud University Nijmegen*
Donders Institute for Brain, Cognition and Behaviour
P.O. Box 9101
6500 HB Nijmegen
The Netherlands

[3] *Karolinska Institutet*
Stockholm Brain Institute
Cognitive Neurophysiology Research Group
17177 Stockholm
Sweden

[4] *Universidade do Algarve*
Institute of Biotechnology & Bioengineering CBME
Cognitive Neuroscience Research Group
8500-139 Faro
Portugal